GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:55:54 ; Search time 29.4726 Seconds
(without alignments)
1123.030 Million cell updates/sec

Title: US-10-734-126-3
Perfect score: 1809
Sequence: 1 MAQKENSYPWPYGRQTAPSG.....PWVRANSRRVLPPSALQSVA 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1507.5	83.3	345	2	JC4665	protein kinase (EC
2	1069	59.1	395	2	JC5975	aurora-related kin
3	1062	58.7	389	2	S52242	protein kinase (EC
4	1050.5	58.1	403	2	JC5974	aurora-related kin
5	1050.5	58.1	407	2	S52243	p46Eg265 protein -
6	940	52.0	294	2	T10690	serine/threonine-s
7	921.5	50.9	282	2	H84653	probable protein k

ALIGNMENTS

RESULT 1
JC4665
protein kinase (EC 2.7.1.37) - mouse
N;Alternate names: serine/threonine protein kinase; STK-1 protein
C;Species: Mus musculus (house mouse)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 17-Mar-1999
C;Accession: JC4665
R;Niwa, H.; Abe, K.; Kunisada, T.; Yamamura, K.
Gene 169, 197-201, 1996
A;Title: Cell-cycle-dependent expression of the STK-1 gene encoding a novel murine putat
A;Reference number: JC4665; MUID:96194801; PMID:8647446
A;Accession: JC4665
A;Molecule type: mRNA
A;Residues: 1-345 <NIW>
A;Cross-references: DDBJ:D21099
A;Experimental source: embryonic stem cell
A;Note: The authors translated the codon GCT for residue 328 as Thr
C;Comment: This enzyme regulates a wide array of cellular responses to changing environme
C;Genetics:
A;Gene: stk-1
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; protein kinase
F;80-332/Domain: protein kinase homology <KIN>
F;88-96/Region: protein kinase ATP-binding motif
F;111/Binding site: ATP (Lys) #status predicted
F;205/Active site: Asp #status predicted

Query Match 83.3%; Score 1507.5; DB 2; Length 345;
Best Local Similarity 83.3%; Pred. No. 3.9e-58;
Matches 289; Conservative 25; Mismatches 24; Indels 9; Gaps 2;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||:|||||| :|: |||:|| |||||||| | ||| |:: || | |||||||: ||
Db          1 MAQKENAYPWPYGSKTSQSGLNTLSQRVLRKEPATTSALALVNWSNSQSTAAPGQKLAEN 60

Qy         61 -------SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQI 113
                     | |: :   : ||||:||||||||||||||||||||| ||||||:||||||
Db         61 KSQGSTASQGSQN--KQPFTIDNFEIGRPLGKGKFGNVYLAREKKSRFIVALKILFKSQI 118

Qy        114 EKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFD 173
              |||||||||||||||||| |||||:|||||||::|||||||||||||||||||| |||
Db        119 EKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPRGELYKELQKSRTFD 178

Qy        174 EQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTM 233
              |||||||||||:||| ||| ||||||||||||||||:|||||||||||||||||||||||
Db        179 EQRTATIMEELSDALTYCHKKKVIHRDIKPENLLLGLQGELKIADFGWSVHAPSLRRKTM 238

Qy        234 CGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFP 293
              |||||||||||||||||| |||||||||||:|||||||| ||:|||||||||||||||
Db        239 CGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHSETYRRIVKVDLKFP 298

Qy        294 ASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              :|||:|||||| |||:||| :|||||:|: |||||||||||||||||
Db        299 SSVPSGAQDLIFKLLKHNPWQRLPLAEVATHPWVRANSRRVLPPSAL 345
```

RESULT 2
JC5975
aurora-related kinase 1 (EC 2.7.-.-) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 16-Aug-2004
C;Accession: JC5975
R;Shindo, M.; Nakano, H.; Kuroyanagi, H.; Shirasawa, T.; Mihara, M.; Gilbert, D.J.; Jenki
Biochem. Biophys. Res. Commun. 244, 285-292, 1998
A;Title: cDNA cloning, expression, subcellulard localization, and chromosomal assignment
A;Reference number: JC5974; MUID:98183439; PMID:9514916
A;Accession: JC5975
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-395 <SHI>
C;Superfamily: protein kinase homology
C;Keywords: phosphotransferase
F;123-374/Domain: protein kinase homology <KIN>

Query Match 59.1%; Score 1069; DB 2; Length 395;
Best Local Similarity 61.4%; Pred. No. 2.2e-39;
Matches 210; Conservative 42; Mismatches 70; Indels 20; Gaps 5;

```
Qy         14 RQTAPSGLSTLPQRVLR----KEPVTPSALVLMS----------RSNVQPTAAPGQKVM 58
              |   ||    :| :  :    ::|  |  |  |           : | || |:
Db         45 RVLCPSNSQRVPSQAQKLGAGQKPAKPKQLPAASVPRPVSRLNNPQKNEQPAASGNDSEK 104

Qy         59 ENSS--GTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKE 116
              | :|   | |   | :|::||:||||||||||||||||::| ||:|||||||:|:||
Db        105 EQASLQKTEDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKA 164

Qy        117 GVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQR 176
              ||||||||:|||:|| |||||||| ||:|  |:|||||||| | :|:||||   |||||
Db        165 NVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQR 224
```

```
Qy        177 TATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGT 236
              ||| : |||:|| ||| |:|||||||||||||  ||||||||||||||| || |||||
Db        225 TATYITELANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGT 284

Qy        237 LDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASV 296
              ||||||||||||||:|||||| :|||||| ||| |||||: :: ||| || :|:  ||  |
Db        285 LDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETY-RISRVEFTFPDFV 343

Qy        297 PTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPS 338
                ||:||||:||:|| |:|| ||:|  |||::||| :  ||:
Db        344 TEGARDLISRLLKHNASQRLTLAEVLEHPWIKANSSK--PPT 383


RESULT 3
S52242
protein kinase (EC 2.7.1.-) p46XlEg22 - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
C;Accession: S52242
R;Roghi, C.; le Guellec, R.; Paris, J.; Couturier, A.; Philippe, M.
submitted to the EMBL Data Library, October 1992
A;Description: Eg2, selected by differential screening encodes a new Xenopus protein kin
A;Reference number: S52242
A;Accession: S52242
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-389 <ROG>
A;Cross-references: EMBL:Z17206; NID:g609279; PIDN:CAA78914.1; PID:g609280
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;119-371/Domain: protein kinase homology <KIN>
F;127-135/Region: protein kinase ATP-binding motif

  Query Match             58.7%;  Score 1062;  DB 2;  Length 389;
  Best Local Similarity   59.8%;  Pred. No. 4.3e-39;
  Matches  207;  Conservative   44;  Mismatches   63;  Indels   32;  Gaps    4;

Qy         20 GLSTLPQRVL--RKEPV----TPSALVLM----------------SRSNVQPTAAPGQKV 57
              | | :||||:   ::||      |:|  |:                    | | |:
Db         32 GPSNVPQRVMMQAQKPVLSNQKPTAQGLLRPATHGHQTSKPQGPNENRNPQQTSHSSTPN 91

Qy         58 MENSSGTPDILT----------RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKV 107
              ||    |   |           : : ::|||||||||||||||||::| ||:||||
Db         92 MEKKGSTDQGKTLAVPKEEGKKKQWCLEDFEIGRPLGKGKFGNVYLARERESKFILALKV 151

Qy        108 LFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQ 167
              |||||:|| |||||||||:|||:|| |||||||| ||:|  |:||||:||| |||::|||
Db        152 LFKSQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDASRVYLILDYAPGGELFRELQ 211

Qy        168 KSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPS 227
              |   ||:||:|  :::||:||:||| ||||||||||||||  ||||||||||||||||
Db        212 KCTRFDDQRSALYIKQLAEALLYCHSKKVIHRDIKPENLLLGSNGELKIADFGWSVHAPS 271

Qy        228 LRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVK 287
              || |:|||||||||||||:|||||| :|||||| ||| ||||: :| ||||||| |
Db        272 SRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFETDTHQETYRRISK 331

Qy        288 VDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRR 333
              |: ::|  |   |:||:||||:|||: ||||  |  |||:  ||::
Db        332 VEFQYPPYVSEEAKDLVSKLLKHNPNHRLPLKGVLEHPWIVKNSQQ 377


RESULT 4
JC5974
aurora-related kinase 1 (EC 2.7.-.-) - human
C;Species: Homo sapiens (man)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 16-Aug-2004
C;Accession: JC5974
R;Shindo, M.; Nakano, H.; Kuroyanagi, H.; Shirasawa, T.; Mihara, M.; Gilbert, D.J.; Jenk
Biochem. Biophys. Res. Commun. 244, 285-292, 1998
A;Title: cDNA cloning, expression, subcellulard localization, and chromosomal assignment
A;Reference number: JC5974; MUID:98183439; PMID:9514916
A;Accession: JC5974
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-403 <SHI>
A;Cross-references: GB:AF008551
C;Superfamily: protein kinase homology
C;Keywords: phosphotransferase
F;131-383/Domain: protein kinase homology <KIN>

  Query Match             58.1%;  Score 1050.5;  DB 2;  Length 403;
  Best Local Similarity   61.7%;  Pred. No. 1.4e-38;
  Matches  209;  Conservative   40;  Mismatches   67;  Indels   23;  Gaps    5;

Qy          4 KENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSSG 63
              :: | | |  |          |       |:|: |||      :| :   |  ||  |:
Db         80 QQTSVPHPVSR--------PLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK- 125

Qy         64 TPDILTRHPTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
                 | : ::|||||||||||||||||||:|  |:|||||||:|:|| |||||||
Db        126 ------RQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLR 179

Qy        124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
              ||:|||:|| |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |
Db        180 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITE 239

Qy        184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
              ||:|| ||| |:|||||||||||||  ||||||||||||||| || |:||||||||||
Db        240 LANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPE 299

Qy        244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
              |||||||:|||||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||
Db        300 MIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDL 359

Qy        304 ISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              ||:||:||||:|  | :|  |||: ||| :   ||  |:
Db        360 ISRLLKHNPSQRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 5
S52243
p46Eg265 protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 07-May-1995 #sequence_revision 03-Aug-1995 #text_change 16-Aug-2004
C;Accession: S52243; S34642; I51695
R;Roghi, C.; le Guellec, R.; Paris, J.; Couturier, A.; Philippe, M.
submitted to the EMBL Data Library, October 1992
A;Description: Eg2, selected by differential screening encodes a new Xenopus protein kina
A;Reference number: S52242
A;Accession: S52243
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-407 <ROG>
A;Cross-references: UNIPROT:Q91820; EMBL:Z17207; NID:g609281; PIDN:CAA78915.1; PID:g60928
R;Bouvet, P.; Omilli, F.; Arlot-Bonnemain, Y.; Legagneux, V.; Roghi, C.; Basset, T.; Osbo
submitted to the EMBL Data Library, June 1993
A;Description: Targetted deadenylation of specific mRNAs in Xenopus embryos by a mechanis
A;Reference number: S34642
A;Accession: S34642
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 328-407 <BOU>
A;Cross-references: EMBL:Z24453; NID:g394756; PIDN:CAA80826.1; PID:g394757
R;Bouvet, P.; Omilli, F.; Arlot-Bonnemains, Y.; Legagneux, V.; Roghi, C.; Bassez, T.; Osb
Mol. Cell. Biol. 14, 1893-1900, 1994
A;Title: The deadenylation conferred by the 3' untranslated region of a developmentally c
A;Reference number: I51695; MUID:94158861; PMID:8114721
A;Accession: I51695
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 328-407 <BO2>
A;Cross-references: EMBL:Z24453; NID:g394756; PIDN:CAA80826.1; PID:g394757
C;Superfamily: protein kinase homology
C;Keywords: ATP; phosphotransferase; protein kinase
F;138-390/Domain: protein kinase homology <KIN>
```

F;146-154/Region: protein kinase ATP-binding motif

Query Match 58.1%; Score 1050.5; DB 2; Length 407;
Best Local Similarity 56.3%; Pred. No. 1.4e-38;
Matches 205; Conservative 46; Mismatches 71; Indels 42; Gaps 4;

```
Qy          11 PYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSR---SNVQPT-------AAPGQKVME- 59
               |   | | | :   ||:|     |     |    :   |: :||       |  | :  :
Db          32 PQSTQFRPPGTAVSAQRILGPSNVPQRVLAQAQKPILSSQKPTTQIPLRPATQGHQSSKP 91

Qy          60 -----------NSSGTPDI------------------LTRHFTIDDFEIGRPLGKGKF 88
                        : | ||::                  : : ::|||||||||||||||
Db          92 QGPNENRNPQQTSHSSTPNVEKKGSTDQGKTSAVPKEEGKKKQWCLEDFEIGRPLGKGKF 151

Qy          89 GNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRR 148
               ||||||||::| ||:|||||||||:|| ||||||||:|||:|| ||||||| ||:|
Db         152 GNVYLARERESKFILALKVLFKSQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDAS 211

Qy         149 RIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLL 208
               |:||||:||| |||::||||   ||:||:|  :::||:||:||| ||||||||||||||
Db         212 RVYLILDYAPGGELFRELQKCTRFDDQRSAMYIKQLAEALLYCHSKKVIHRDIKPENLLL 271

Qy         209 GLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLV 268
               |  |||||||||||||||| || |:||||||||||||||||:| |||| :|||||| ||
Db         272 GSNGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDETVDLWSLGVLCYEFLV 331

Qy         269 GNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVR 328
               | ||||: :| |||||| ||: ::|  |    :|||:||||:|||: ||||  |  |||:
Db         332 GKPPFETDTHQETYRRISKVEFQYPPYVSEEARDLVSKLLKHNPNHRLPLKGVLEHPWII 391

Qy         329 ANSR 332
                 ||:
Db         392 KNSQ 395
```

RESULT 6
T10690
serine/threonine-specific protein kinase homolog T16I18.40 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10690
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16533
A;Accession: T10690
A;Molecule type: DNA
A;Residues: 1-294 <BEV>
A;Cross-references: UNIPROT:Q9M077; EMBL:AL049915; GSPDB:GN00062; ATSP:T16I18.40
A;Experimental source: cultivar Columbia; BAC clone T16I18
C;Genetics:
A;Gene: ATSP:T16I18.40
A;Map position: 4
A;Introns: 12/3; 52/3; 101/3; 133/3; 165/3; 198/1; 237/2; 261/3
C;Superfamily: kinase-related transforming protein; protein kinase homology
F;29-282/Domain: protein kinase homology <KIN>

Query Match 52.0%; Score 940; DB 2; Length 294;
Best Local Similarity 60.3%; Pred. No. 5.5e-34;
Matches 179; Conservative 47; Mismatches 63; Indels 8; Gaps 4;

```
Qy          49 PTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVL 108
               ||    |:  | |  :     : :|: ||:||:|||:||||:||||||||:|: :||||||
Db           4 PTETQHQE-KEASDASAAAAQKRWTLSDFDIGKPLGRGKFGHVYLAREKRSNHVVALKVL 62

Qy         109 FKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQK 168
               ||||:::  |||||||:|||:|| |||||||| ||||::|:|||||||| ||||||:|||
Db          63 FKSQLQQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQKRVYLILEYAARGELYKDLQK 122

Qy         169 SCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSL 228
                  | |:| || :  || ||:||||| |||||||||||:| :||||||||||||  :
Db         123 CKYFSERRAATYVASLARALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN- 181

Qy         229 RRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKV 288
               ||:|||||||||||||:|   |:  ||:|  :|:|||| | |:||||:  |::||||||:|
Db         182 RRRTMCGTLDYLPPEMVESVEHDASVDIWSLGILCYEFLYGVPPFEAMEHSDTYRRIVQV 241

Qy         289 DLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
               |||||    :   |:||||::|    |:|||| ::  |||:  |:      || :  |
Db         242 DLKFPPKPIISASAKDLISQMLVKESSQRLPLHKLLEHPWIVQNA----DPSGIYRV 294
```

RESULT 7
H84653
probable protein kinase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84653
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84653
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-282 <STO>
A;Cross-references: UNIPROT:O82309; GB:AE002093; NID:g3643610; PIDN:AAC42257.1; GSPDB:GN
C;Genetics:
A;Gene: At2g25880
A;Map position: 2
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 50.9%; Score 921.5; DB 2; Length 282;
Best Local Similarity 63.7%; Pred. No. 3.3e-33;
Matches 174; Conservative 42; Mismatches 50; Indels 7; Gaps 3;

```
Qy          70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
               : :|  ||:||:|||:||||:|||||||:|  ||||||||||:|:::  ||||||||:|||
Db          12 KRWTTSDFDIGKPLGRGKFGHVYLAREKRSDHIVALKVLFKAQLQQSQVEHQLRREVEIQ 71

Qy         130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
               :|| |||||||| ||||::|:|||||| |||||||||    | |:| || :  || ||:
Db          72 SHLRHPNILRLYGYFYDQKRVYLILEYAVRGELYKELQKCKYFSERRAATYVASLARALI 131

Qy         190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
               ||||| |||||||||||:| :||||||||||||  : ||:||||||||||||:|
Db         132 YCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 190

Qy         250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
               |:  ||:| :|:|||| | | ||||:  |:|||:|||:||||||    | : |:||||::
Db         191 HDASVDIWSLGILCYEFLYGVPPFEAREHSETYKRIVQVDLKFPPKPIVSSSAKDLISQM 250

Qy         308 LRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
               |    ::|| | ::  |||:  |:     || |
Db         251 LVKESTQRLALHKLLEHPWIVQNA----DPSGL 279
```

Search completed: June 20, 2005, 20:07:57
Job time : 32.4726 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:59:50 ; Search time 44.2383 Seconds
(without alignments)
680.035 Million cell updates/sec

Title: US-10-734-126-4
Perfect score: 2110
Sequence: 1 MDRSKENCISGPVKATAPVG.....TANSSKPSNCQNKESASKQS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 13

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2110	100.0	403	2	US-08-755-728-4	Sequence 4, Appli
2	2110	100.0	403	2	US-08-974-655-4	Sequence 4, Appli
3	2110	100.0	403	3	US-09-283-011-4	Sequence 4, Appli
4	2110	100.0	403	4	US-09-012-135A-4	Sequence 4, Appli
5	2110	100.0	488	4	US-09-949-016-10461	Sequence 10461, A
6	2110	100.0	488	4	US-09-949-016-10462	Sequence 10462, A
7	1058	50.1	363	4	US-09-949-016-9826	Sequence 9826, Ap
8	1057.5	50.1	344	2	US-08-755-728-3	Sequence 3, Appli
9	1057.5	50.1	344	2	US-08-974-655-3	Sequence 3, Appli
10	1057.5	50.1	344	3	US-09-283-011-3	Sequence 3, Appli
11	1057.5	50.1	344	4	US-09-012-135A-3	Sequence 3, Appli
12	1050	49.8	343	4	US-09-485-534-4	Sequence 4, Appli
13	1046	49.6	347	2	US-09-016-000-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-755-728-4
; Sequence 4, Application US/08755728
; Patent No. 5962312
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Mossie, Kevin
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
; TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,728
; FILING DATE: No. 5962312ember 25, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,809
; FILING DATE: December 18, 1995
; APPLICATION NUMBER: 60/023,943
; FILING DATE: August 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 223/113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 403 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-755-728-4

Query Match 100.0%; Score 2110; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.6e-179;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 2
US-08-974-655-4
; Sequence 4, Application US/08974655
; Patent No. 5972676
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,655
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 5972676ember 25, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  223/113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-974-655-4

  Query Match             100.0%;  Score 2110;  DB 2;  Length 403;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-179;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 3
US-09-283-011-4
; Sequence 4, Application US/09283011
; Patent No. 6207401
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/283,011
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/012,135
;      FILING DATE:  January 22, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 6207401ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-283-011-4
```

```
  Query Match           100.0%;  Score 2110;  DB 3;  Length 403;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-179;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 4
US-09-012-135A-4
; Sequence 4, Application US/09012135A
; Patent No. 6716575
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,135A
;      FILING DATE:  January 22, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/005,268
;      FILING DATE:  January 9, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 6716575ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-012-135A-4

  Query Match           100.0%;  Score 2110;  DB 4;  Length 403;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-179;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 5
US-09-949-016-10461
; Sequence 10461, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10461
;   LENGTH: 488
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10461
```

```
  Query Match             100.0%;  Score 2110;  DB 4;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-179;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 145

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 205

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 265

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 325

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 385

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 445

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        446 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 488


RESULT 6
US-09-949-016-10462
; Sequence 10462, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10462
;   LENGTH: 488
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10462

  Query Match             100.0%;  Score 2110;  DB 4;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-179;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 145

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 205

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 265

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 325

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 385

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 445

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        446 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 488


RESULT 7
US-09-949-016-9826
; Sequence 9826, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9826
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-9826

  Query Match             50.1%;  Score 1058;  DB 4;  Length 363;
  Best Local Similarity   59.2%;  Pred. No. 6e-86;
  Matches  218;  Conservative   43;  Mismatches   83;  Indels   24;  Gaps    6;

Qy         43 QAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPL 102
              ::   | ||:         ||| |   |   |     :| :|  | |        |:|:
Db          3 ESSSALDPSSPPPFSLRMAQKENSYPWPYGRQTAPSGLST-LPQRVLR--------KEPV 53

Qy        103 -PSA----PENNPEEELASKQKNEESKK-------RQWALEDFEIGRPLGKGKFGNVYLA 150
               |||      :| :   |  ||  |:         | : ::||||||||||||||||||
Db         54 TPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLA 113

Qy        151 REKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLIL 210
              |||:| ||:|||||||:|:|| |||||||||:|||:|| |||||||  ||:|   |:||||
Db        114 REKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLIL 173

Qy        211 EYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGEL 270
              |||| | :|:|||||    |||||||| : |||:|| ||| |:|||||||||||||  |||
Db        174 EYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGEL 233

Qy        271 KIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFE 330
              |||||||||||| || |:|||||||||||||||:|||||| :|||||| ||| ||||
Db        234 KIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFE 293

Qy        331 ANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK- 389
              : :: |||:|| :|:  ||  |  ||:||||:||:||||:|  | :|  |||: ||| :
Db        294 SASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRV 353

Qy        390 --PSNCQN 395
```

```
                 ||  |:
Db        354 LPPSALQS 361


RESULT 8
US-08-755-728-3
; Sequence 3, Application US/08755728
; Patent No. 5962312
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  PastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/755,728
;      FILING DATE:  No. 5962312ember 25, 1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  223/113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  344 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-755-728-3

  Query Match           50.1%;  Score 1057.5;  DB 2;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 6.2e-86;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::||||||||||||||||||||:| ||:||||||:|:|| ||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
                |:|||||||||||||| ||||||||||||||| || |:||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 9
US-08-974-655-3
; Sequence 3, Application US/08974655
; Patent No. 5972676
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,655
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 5972676ember 25, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  223/113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  344 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-974-655-3

  Query Match           50.1%;  Score 1057.5;  DB 2;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 6.2e-86;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
```

```
Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 10
US-09-283-011-3
; Sequence 3, Application US/09283011
; Patent No. 6207401
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/283,011
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/012,135
;      FILING DATE:  January 22, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 6207401ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  344 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-283-011-3

  Query Match           50.1%;  Score 1057.5;  DB 3;  Length 344;
  Best Local Similarity  63.2%;  Pred. No. 6.2e-86;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 11
US-09-012-135A-3
; Sequence 3, Application US/09012135A
; Patent No. 6716575
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,135A
;      FILING DATE:  January 22, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/005,268
;      FILING DATE:  January 9, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 6716575ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
```

```
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  344 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-012-135A-3

  Query Match           50.1%;  Score 1057.5;  DB 4;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 6.2e-86;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:        | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDPEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDPEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:|||    |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||:|||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 12
US-09-485-534-4
; Sequence 4, Application US/09485534
; Patent No. 6759212
; GENERAL INFORMATION:
;  APPLICANT:  TATSUKA, MASAAKI
;  APPLICANT:  TERADA, YASUHIKO
;  TITLE OF INVENTION:  CELL CYCLE-REGULATING PROTEINS
;  FILE REFERENCE:  050499/0102
;  CURRENT APPLICATION NUMBER: US/09/485,534
;  CURRENT FILING DATE:  2000-02-14
;  PRIOR APPLICATION NUMBER:  PCT/JP98/03641
;  PRIOR FILING DATE:  1998-08-17
;  PRIOR APPLICATION NUMBER:  235371/1997
;  PRIOR FILING DATE:  1997-08-15
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Rattus sp.
US-09-485-534-4

  Query Match           49.8%;  Score 1050;  DB 4;  Length 343;
  Best Local Similarity   57.6%;  Pred. No. 2.9e-85;
  Matches  204;  Conservative   51;  Mismatches   73;  Indels   26;  Gaps    4;

Qy         36 PLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNT 95
              | |  | :|  |  :   ||| |: :  |:  : : |:                 :|:
Db          9 PWPYGSKTSQSGL--NTLPQRV-LRKEPAVTPAQALMNR-----------------SNS 47

Qy         96 QKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQS 155
              | :  |     ||     |  |      :: : :::||||||||||||||||||||:|
Db         48 QSTAVPGQKLTENKGATALQGSQ-----SRQPFTIDNFEIGRPLGKGKFGNVYLAREKKS 102

Qy        156 KFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPL 215
              :||:|||:|||:|:|| ||||||||||:|||:||:|||||:|| ||:|  |:||||||
Db        103 RFIVALKILFKSQIEKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPR 162

Qy        216 GTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADF 275
              | :|:|||    |||||||| : ||::|| ||| |:|||||||||||||  ||||||||
Db        163 GELYKELQKSGTFDEQRTATIMEELSDALMYCHKKKVIHRDIKPENLLLGLQGELKIADF 222

Qy        276 GWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQ 335
              |||||||| || |:|||||||||||||||:| |||| :||||| :|| ||||: ::
Db        223 GWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHS 282

Qy        336 ETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
              |||:|| :|:  ||  :  ||:||||:||||||||  | :|  |||: ||| :
Db        283 ETYRRIVKVDLKFPSSMPLGAKDLISKLLKHNPSQRLPLEQVSAHPWVRANSRR 336


RESULT 13
US-09-016-000-1
; Sequence 1, Application US/09016000
; Patent No. 5962232
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    TITLE OF INVENTION:  PROTEIN KINASE MOLECULES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,000
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0465 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  347 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
```

```
;      LIBRARY:  HMC1NOT01
;      CLONE:  2940
US-09-016-000-1

  Query Match             49.6%;  Score 1046;  DB 2;  Length 347;
  Best Local Similarity   62.7%;  Pred. No. 6.6e-85;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   26;  Gaps    6;

Qy          82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
               :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db          22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy         130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
               ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db          74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy         190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTAT---YITELANALSY 246
               |||||||| ||:|  |:|||||||| | :|:|||||   ||||||||    : |||:|| |
Db         134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATVRAIMEELADALMY 193

Qy         247 CHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMH 306
               || |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||
Db         194 CHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMH 253

Qy         307 DEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKH 366
               :|||||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||:|
Db         254 NEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRH 313

Qy         367 NPSQRPMLREVLEHPWITANSSK---PSNCQN 395
               |||:|  | :|  |||: ||| :   ||  |:
Db         314 NPSERLPLAQVSAHPWVRANSRRVLPPSALQS 345


Search completed: June 20, 2005, 20:13:51
Job time : 47.2383 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:47:43 ; Search time 77.6867 Seconds
(without alignments)
1991.894 Million cell updates/sec

Title: US-10-734-126-4
Perfect score: 2110
Sequence: 1 MDRSKENCISGPVKATAPVG.....TANSSKPSNCQNKESASKQS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 37

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2110	100.0	403	9	US-09-012-135A-4	Sequence 4, Appli
2	2110	100.0	403	11	US-09-784-332-4	Sequence 4, Appli
3	2110	100.0	403	15	US-10-087-684-47	Sequence 47, Appl
4	2110	100.0	403	15	US-10-218-779-47	Sequence 47, Appl
5	2110	100.0	403	16	US-10-734-126-4	Sequence 4, Appli
6	2110	100.0	420	15	US-10-264-049-3090	Sequence 3090, Ap
7	2109	100.0	403	14	US-10-060-065-12	Sequence 12, Appl
8	2109	100.0	403	14	US-10-059-585-33	Sequence 33, Appl
9	2109	100.0	403	15	US-10-087-684-48	Sequence 48, Appl
10	2109	100.0	403	15	US-10-218-779-48	Sequence 48, Appl
11	2104	99.7	403	15	US-10-295-027-1206	Sequence 1206, Ap
12	2104	99.7	403	15	US-10-087-684-49	Sequence 49, Appl
13	2104	99.7	403	15	US-10-218-779-49	Sequence 49, Appl
14	2104	99.7	403	15	US-10-188-832-35	Sequence 35, Appl
15	2104	99.7	403	16	US-10-751-736-65	Sequence 65, Appl
16	2104	99.7	403	17	US-10-781-581-187	Sequence 187, App
17	2103	99.7	403	14	US-10-209-324-2	Sequence 2, Appli
18	2095	99.3	403	14	US-10-026-021-7	Sequence 7, Appli
19	2095	99.3	403	15	US-10-087-684-50	Sequence 50, Appl
20	2095	99.3	403	15	US-10-218-779-50	Sequence 50, Appl
21	1935.5	91.7	402	15	US-10-087-684-51	Sequence 51, Appl
22	1935.5	91.7	402	15	US-10-218-779-51	Sequence 51, Appl
23	1924	91.2	403	15	US-10-087-684-14	Sequence 14, Appl
24	1924	91.2	403	15	US-10-218-779-14	Sequence 14, Appl
25	1690	80.1	322	16	US-10-664-421-125	Sequence 125, App
26	1057.5	50.1	344	9	US-09-012-135A-3	Sequence 3, Appli
27	1057.5	50.1	344	11	US-09-784-332-3	Sequence 3, Appli
28	1057.5	50.1	344	14	US-10-060-065-13	Sequence 13, Appl
29	1057.5	50.1	344	14	US-10-059-585-34	Sequence 34, Appl
30	1057.5	50.1	344	14	US-10-171-311-214	Sequence 214, App
31	1057.5	50.1	344	15	US-10-295-027-203	Sequence 203, App
32	1057.5	50.1	344	15	US-10-173-999-101	Sequence 101, App
33	1057.5	50.1	344	15	US-10-188-832-93	Sequence 93, Appl
34	1057.5	50.1	344	16	US-10-734-126-3	Sequence 3, Appli
35	1050	49.8	343	15	US-10-429-849-4	Sequence 4, Appli
36	1046	49.6	347	9	US-09-974-298-136	Sequence 136, App
37	1045	49.5	320	15	US-10-403-571-98	Sequence 98, Appl

ALIGNMENTS

```
RESULT 1
US-09-012-135A-4
; Sequence 4, Application US/09012135A
; Patent No. US20020081578A1
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,135A
;      FILING DATE:  January 22, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/005,268
;      FILING DATE:  January 9, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. US20020081578A1ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (213) 489-1600
;       TELEFAX:  (213) 955-0440
;       TELEX:  67-3510
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  403 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-09-012-135A-4

  Query Match             100.0%;  Score 2110;  DB 9;  Length 403;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-147;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 2
US-09-784-332-4
; Sequence 4, Application US/09784332
; Publication No. US20050002938A1
;   GENERAL INFORMATION:
;       APPLICANT: Plowman, Gregory
;                  Mossie, Kevin
;       TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
;                           AND/OR AUR-2 RELATED DISORDERS
;       NUMBER OF SEQUENCES: 39
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street
;                    Suite 4700
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: U.S.A.
;            ZIP: 90071-2066
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                         storage
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: IBM P.C. DOS 5.0
;            SOFTWARE: FastSEQ for Windows 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/784,332
;            FILING DATE: 16-Feb-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/283,011
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 09/012,135
;            FILING DATE: January 22, 1998
;            APPLICATION NUMBER: 08/755,728
;            FILING DATE: November 25, 1996
;            APPLICATION NUMBER: 60/023,943
;            FILING DATE: August 14, 1996
;            APPLICATION NUMBER: 60/008,809
;            FILING DATE: December 18, 1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Warburg, Richard J.
;            REGISTRATION NUMBER: 32,327
;            REFERENCE/DOCKET NUMBER: 231/282
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (213) 489-1600
;            TELEFAX: (213) 955-0440
;            TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 403 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-784-332-4

  Query Match             100.0%;  Score 2110;  DB 11;  Length 403;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-147;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 3
US-10-087-684-47
; Sequence 47, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
```

```
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 47
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-47

  Query Match             100.0%;  Score 2110;  DB 15;  Length 403;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-147;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 4
US-10-218-779-47
; Sequence 47, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-47

  Query Match             100.0%;  Score 2110;  DB 15;  Length 403;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-147;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 5
US-10-734-126-4
; Sequence 4, Application US/10734126
; Publication No. US20040265852A1
;   GENERAL INFORMATION:
;        APPLICANT: Plowman, Gregory
;                   Mossie, Kevin
;        TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
;                            AND/OR AUR-2 RELATED DISORDERS
;        NUMBER OF SEQUENCES: 39
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/734,126
;             FILING DATE: 15-Dec-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/784,332
;             FILING DATE: 16-Feb-2001
;             APPLICATION NUMBER: 09/283,011
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 09/012,135
;             FILING DATE: January 22, 1998
;             APPLICATION NUMBER: 08/755,728
;             FILING DATE: November 25, 1996
;             APPLICATION NUMBER: 60/023,943
;             FILING DATE: August 14, 1996
;             APPLICATION NUMBER: 60/008,809
;             FILING DATE: December 18, 1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Warburg, Richard J.
;             REGISTRATION NUMBER: 32,327
;             REFERENCE/DOCKET NUMBER: 231/282
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 403 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-734-126-4

  Query Match           100.0%;  Score 2110;  DB 16;  Length 403;
  Best Local Similarity  100.0%;  Pred. No. 3.2e-147;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 6
US-10-264-049-3090
; Sequence 3090, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3090
;   LENGTH: 420
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-049-3090

  Query Match           100.0%;  Score 2110;  DB 15;  Length 420;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-147;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 77

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        78 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 137

Qy       121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       138 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 197

Qy       181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       198 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 257

Qy       241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       258 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 317

Qy       301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       318 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 377

Qy       361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
            |||||||||||||||||||||||||||||||||||||||||||
Db       378 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 420


RESULT 7
US-10-060-065-12
; Sequence 12, Application US/10060065
; Publication No. US20030017480A1
; GENERAL INFORMATION:
;  APPLICANT: Toshio Ota
;  APPLICANT:  Takao Isogai
;  APPLICANT:  Tetsuo Nishikawa
;  APPLICANT:  Koji Hayashi
;  APPLICANT:  Kaoru Otsuka
;  APPLICANT:  Jun-Ichi Yamamoto
;  APPLICANT:  Shizuko Ishii
;  APPLICANT:  Tomoyasu Sugiyama
;  APPLICANT:  Ai Wakamatsu
;  APPLICANT:  Keiichi Nagai
;  APPLICANT:  Tetsuji Otsuki
;  APPLICANT:  Shin-Ichi Funahashi
;  APPLICANT:  Chiaki Senoo
;  APPLICANT:  Jun-Ichi Nezu
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
;  FILE REFERENCE: 06501-099002
;  CURRENT APPLICATION NUMBER: US/10/060,065
;  CURRENT FILING DATE:  2002-01-29
;  PRIOR APPLICATION NUMBER: PCT/JP00/05061
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: US 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: US 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: JP 11-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-060-065-12

  Query Match           100.0%;  Score 2109;  DB 14;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 3.8e-147;
  Matches  402;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db         1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy        61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy       121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy       181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy       241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy       301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy       361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
            |||||||||||||||||||||||||||||||||||||||||||
Db       361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 8
US-10-059-585-33
; Sequence 33, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
;  APPLICANT: Ota, Toshio
;  APPLICANT:  Isogai, Takao
;  APPLICANT:  Nishikawa, Tetsuo
;  APPLICANT:  Hayashi, Koji
;  APPLICANT:  Otsuka, Kaoru
;  APPLICANT:  Yamamoto, Jun-ichi
;  APPLICANT:  Ishii, Shizuko
;  APPLICANT:  Sugiyama, Tomoyasu
;  APPLICANT:  Wakamatsu, Ai
;  APPLICANT:  Nagai, Keiichi
;  APPLICANT:  Otsuki, Tetsuji
;  APPLICANT:  Funahashi, Shin-Ichi
;  APPLICANT:  Senoo, Chiaki
;  APPLICANT:  Nezu, Jun-Ichi
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
;  TITLE OF INVENTION:  KINASE/PROTEIN PHOSPHATASE
;  FILE REFERENCE: 06501-098001
;  CURRENT APPLICATION NUMBER: US/10/059,585
;  CURRENT FILING DATE:  2002-01-29
;  PRIOR APPLICATION NUMBER: PCT/JP00/05060
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: US 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: US 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 11-248036
;  PRIOR FILING DATE: 1999-07-29
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-059-585-33

  Query Match           100.0%;  Score 2109;  DB 14;  Length 403;
```

```
  Best Local Similarity   99.8%;  Pred. No. 3.8e-147;
  Matches  402;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 9
US-10-087-684-48
; Sequence 48, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 48
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-48

  Query Match           100.0%;  Score 2109;  DB 15;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 3.8e-147;
  Matches  402;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 10
US-10-218-779-48
; Sequence 48, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
```

```
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-48

  Query Match           100.0%;  Score 2109;  DB 15;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 3.8e-147;
  Matches  402;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 11
US-10-295-027-1206
; Sequence 1206, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1206
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-295-027-1206

  Query Match           99.7%;  Score 2104;  DB 15;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 8.9e-147;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 12
US-10-087-684-49
; Sequence 49, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 49
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-49

  Query Match            99.7%;  Score 2104;  DB 15;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 8.9e-147;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 13
US-10-218-779-49
; Sequence 49, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-49

  Query Match            99.7%;  Score 2104;  DB 15;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 8.9e-147;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
```

```
Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 14
US-10-188-832-35
; Sequence 35, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-35

Query Match 99.7%; Score 2104; DB 15; Length 403;
Best Local Similarity 99.8%; Pred. No. 8.9e-147;
Matches 402; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 15
US-10-751-736-65
; Sequence 65, Application US/10751736
; Publication No. US20040265230A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Martinez, Robert
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM100927 (031896-002000)
; CURRENT APPLICATION NUMBER: US/10/751,736
; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
; PRIOR FILING DATE: 2003-01-06
; NUMBER OF SEQ ID NOS: 54873
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 65
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-751-736-65

Query Match 99.7%; Score 2104; DB 16; Length 403;
Best Local Similarity 99.8%; Pred. No. 8.9e-147;
Matches 402; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 16
US-10-781-581-187

```
; Sequence 187, Application US/10781581
; Publication No. US20050019746A1
; GENERAL INFORMATION:
;  APPLICANT: EiRx Therapeutics Ltd.
;  APPLICANT:  Seery, Liam
;  APPLICANT:  Hayes, Ian
;  APPLICANT:  Murphy, Finbarr
;  TITLE OF INVENTION: Apoptosis-Related Kinase/GPCRs
;  FILE REFERENCE: 8912/2015
;  CURRENT APPLICATION NUMBER: US/10/781,581
;  CURRENT FILING DATE:  2004-02-18
;  PRIOR APPLICATION NUMBER: US 10/764,238
;  PRIOR FILING DATE: 2004-01-23
;  PRIOR APPLICATION NUMBER: US 60/457,533
;  PRIOR FILING DATE: 2003-03-25
;  PRIOR APPLICATION NUMBER: UK 0301566.6
;  PRIOR FILING DATE: 2003-01-23
;  NUMBER OF SEQ ID NOS: 226
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 187
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-781-581-187

  Query Match             99.7%;  Score 2104;  DB 17;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 8.9e-147;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 17
US-10-209-324-2
; Sequence 2, Application US/10209324
; Publication No. US20030108910A1
; GENERAL INFORMATION:
;  APPLICANT: UNIVERSITY OF CALIFORNIA SAN FRANCISCO
;  APPLICANT:  TOLAND, Amanda E.
;  APPLICANT:  BALMAIN, Allan
;  TITLE OF INVENTION: STK15 (STK6) GENE POLYMORPHISM AND METHODS OF DETERMINING CANCER
;  FILE REFERENCE: UCSF1120-2
;  CURRENT APPLICATION NUMBER: US/10/209,324
;  CURRENT FILING DATE:  2002-07-29
;  PRIOR APPLICATION NUMBER: US 60/334,145
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: US 60/308,911
;  PRIOR FILING DATE: 2001-07-27
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: Xaa is Ile or Phe
US-10-209-324-2

  Query Match             99.7%;  Score 2103;  DB 14;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-146;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQXPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 18
US-10-026-021-7
; Sequence 7, Application US/10026021
; Publication No. US20030027756A1
; GENERAL INFORMATION:
;  APPLICANT: Hitoshi, Yasumichi
;  APPLICANT:  Demo, Susan
;  APPLICANT:  Jenkins, Yonchu
;  APPLICANT:  Rigel Pharmaceuticals, Inc.
;  TITLE OF INVENTION: SAK:  Modulation of Cellular Proliferation for
;  TITLE OF INVENTION:  Treatment of Cancer
;  FILE REFERENCE: 021044-001210US
;  CURRENT APPLICATION NUMBER: US/10/026,021
;  CURRENT FILING DATE:  2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/309,632
;  PRIOR FILING DATE: 2001-08-01
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human ARK mitotic kinase
US-10-026-021-7

  Query Match             99.3%;  Score 2095;  DB 14;  Length 403;
```

```
  Best Local Similarity   99.3%;  Pred. No. 4.1e-146;
  Matches  400;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQQTSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 19
US-10-087-684-50
; Sequence 50, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 50
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-50

  Query Match             99.3%;  Score 2095;  DB 15;  Length 403;
  Best Local Similarity   99.3%;  Pred. No. 4.1e-146;
  Matches  400;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQQTSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 20
US-10-218-779-50
; Sequence 50, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
```

```
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-50

  Query Match             99.3%;  Score 2095;  DB 15;  Length 403;
  Best Local Similarity   99.3%;  Pred. No. 4.1e-146;
  Matches  400;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQQTSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 21
US-10-087-684-51
; Sequence 51, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 51
;   LENGTH: 402
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-51

  Query Match             91.7%;  Score 1935.5;  DB 15;  Length 402;
  Best Local Similarity   92.9%;  Pred. No. 2.3e-134;
  Matches  378;  Conservative    3;  Mismatches   17;  Indels    9;  Gaps    3;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPS----APENNPEEELAS 116
              ||||||||||||||||||||||||||||||||||||||||||||    |  :
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSHLKIILRRNWHQN--R 118

Qy        117 KQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEH 176
              | ||:   |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 KMKNQ---KEAVALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEH 175

Qy        177 QLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATY 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 QLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTANL 235

Qy        237 ITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYL 296
                :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 YNRIANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYL 295

Qy        297 PPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGA 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 PPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGA 355

Qy        357 RDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        356 RDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 402


RESULT 22
```

```
US-10-218-779-51
; Sequence 51, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
;   LENGTH: 402
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-51

  Query Match             91.7%;  Score 1935.5;  DB 15;  Length 402;
  Best Local Similarity   92.9%;  Pred. No. 2.3e-134;
  Matches  378;  Conservative    3;  Mismatches   17;  Indels    9;  Gaps    3;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPS----APENNPEEELAS 116
              ||||||||||||||||||||||||||||||||||||||||||||    |  :
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSHLKIILRRNWHQN--R 118

Qy        117 KQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEH 176
              | ||:   |  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 KMKNQ---KEAVALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEH 175

Qy        177 QLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATY 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 QLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTANL 235

Qy        237 ITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYL 296
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 YNRIANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYL 295

Qy        297 PPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGA 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 PPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGA 355

Qy        357 RDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        356 RDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 402


RESULT 23
US-10-087-684-14
; Sequence 14, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 14
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-14

  Query Match             91.2%;  Score 1924;  DB 15;  Length 403;
  Best Local Similarity   91.8%;  Pred. No. 1.6e-133;
  Matches  370;  Conservative   11;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||:|:|||||||:|  |||||: |||||||||||||| |||||| ||||||||||||||
Db          1 MDQSEENCISGPVEAKTPVGGPEHVLVTQQFPCQNPLPANSGQAQWVLCPSNSSQRVPLQ 60
```

```
Qy          61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
               ||||||||||||||||||||||||||||| |||||||| |||   ||||||||||||||||||
Db          61 AQKLVSSHKPVQNQKQKQLQATSVPHPASRPLNNTQNSKQSPLSAPENNPEEELASKQKN 120

Qy         121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
               |||||||||||| ||||| |||||||||||||||||||:||||||||||||||||||||
Db         121 EESKKRQWALEDLEIGRPPGKGKFGNVYLAREKQSKFILALRVLFKAQLEKAGVEHQLRR 180

Qy         181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
               ||||||||:||||:||||||||||||||||| || ||  ||||||||||||||||||||
Db         181 EVEIQSHLQHPNIIRLYGYFHDATRVYLILEYTPLETVNTELQKLSKFDEQRTATYITEL 240

Qy         241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
               |:||||||||  |||||||||||||||||||:||:||||  ||||||||||||||||||||
Db         241 ASALSYCHSKTVIHRDIKPENLLLGSAGELEIANFGWSEHAPSSRRTTLCGTLDYLPPEM 300

Qy         301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
               |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db         301 IEGRMHDEKVDLWSLGVLCCEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy         361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
               ||||||  ||||||||||||:|||||||||||||||||| || |
Db         361 SRLLKHVPSQRPMLREVLEYPWITANSSKPSNCQNKESTSKYS 403
```

RESULT 24
US-10-218-779-14
; Sequence 14, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patturajan, Meera
; APPLICANT: Shenoy, Suresh
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Malyankar, Uriel
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles
; APPLICANT: Gangolli, Esha
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-214
; CURRENT APPLICATION NUMBER: US/10/218,779
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,-926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-779-14

Query Match 91.2%; Score 1924; DB 15; Length 403;
Best Local Similarity 91.8%; Pred. No. 1.6e-133;
Matches 370; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

```
Qy           1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
               ||:|:|||||||||:|  |||||: |||||||||||||| |||||| ||||||||||||||
Db           1 MDQSEENCISGPVEAKTPVGGPEHVLVTQQFPCQNPLPANSGQAQWVLCPSNSSQRVPLQ 60

Qy          61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
               |||||||||||||||||||||||||||| |||||||| |||   ||||||||||||||||||
Db          61 AQKLVSSHKPVQNQKQKQLQATSVPHPASRPLNNTQNSKQSPLSAPENNPEEELASKQKN 120

Qy         121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
               |||||||||||| ||||| |||||||||||||||||||:||||||||||||||||||||
Db         121 EESKKRQWALEDLEIGRPPGKGKFGNVYLAREKQSKFILALRVLFKAQLEKAGVEHQLRR 180

Qy         181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
               ||||||||:||||:||||||||||||||||| || ||  ||||||||||||||||||||
Db         181 EVEIQSHLQHPNIIRLYGYFHDATRVYLILEYTPLETVNTELQKLSKFDEQRTATYITEL 240

Qy         241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
               |:||||||||  |||||||||||||||||||:||:||||  ||||||||||||||||||||
Db         241 ASALSYCHSKTVIHRDIKPENLLLGSAGELEIANFGWSEHAPSSRRTTLCGTLDYLPPEM 300

Qy         301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
               |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db         301 IEGRMHDEKVDLWSLGVLCCEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy         361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
               ||||||  ||||||||||||:|||||||||||||||||| || |
Db         361 SRLLKHVPSQRPMLREVLEYPWITANSSKPSNCQNKESTSKYS 403
```

RESULT 25
US-10-664-421-125
; Sequence 125, Application US/10664421
; Publication No. US20040142864A1
; GENERAL INFORMATION:
; APPLICANT: BREMER, RYAN
; APPLICANT: IBRAHIM, PRABHA
; APPLICANT: KUMAR, ABHINAV
; APPLICANT: MANDIYAN, VALSAN
; APPLICANT: MILBURN, MICHAEL V.
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE
; FILE REFERENCE: 039363/0703
; CURRENT APPLICATION NUMBER: US/10/664,421
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/412,341
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/411,398
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 125
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-664-421-125

Query Match 80.1%; Score 1690; DB 16; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.2e-116;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          82 TSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGK 141
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 TSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGK 60
```

```
Qy        142 GKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFH 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFH 120

Qy        202 DATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPEN 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPEN 180

Qy        262 LLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYE 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYE 240

Qy        322 FLVGKPPPEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHP 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 FLVGKPPPEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHP 300

Qy        382 WITANSSKPSNCQNKESASKQS 403
              ||||||||||||||||||||||
Db        301 WITANSSKPSNCQNKESASKQS 322


RESULT 26
US-09-012-135A-3
; Sequence 3, Application US/09012135A
; Patent No. US20020081578A1
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,135A
;      FILING DATE:  January 22, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/005,268
;      FILING DATE:  January 9, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. US20020081578A1ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  344 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-012-135A-3

  Query Match           50.1%;  Score 1057.5;  DB 9;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 27
US-09-784-332-3
; Sequence 3, Application US/09784332
; Publication No. US20050002938A1
;   GENERAL INFORMATION:
;        APPLICANT: Plowman, Gregory
;                   Mossie, Kevin
;        TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
;                            AND/OR AUR-2 RELATED DISORDERS
;        NUMBER OF SEQUENCES: 39
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/784,332
;             FILING DATE: 16-Feb-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/283,011
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 09/012,135
;             FILING DATE: January 22, 1998
;             APPLICATION NUMBER: 08/755,728
;             FILING DATE: November 25, 1996
;             APPLICATION NUMBER: 60/023,943
;             FILING DATE: August 14, 1996
;             APPLICATION NUMBER: 60/008,809
;             FILING DATE: December 18, 1995
```

```
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Warburg, Richard J.
;            REGISTRATION NUMBER: 32,327
;            REFERENCE/DOCKET NUMBER: 231/282
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (213) 489-1600
;            TELEFAX: (213) 955-0440
;            TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 344 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-784-332-3

  Query Match            50.1%;  Score 1057.5;  DB 11;  Length 344;
  Best Local Similarity  63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 28
US-10-060-065-13
; Sequence 13, Application US/10060065
; Publication No. US20030017480A1
; GENERAL INFORMATION:
;  APPLICANT: Toshio Ota
;  APPLICANT:  Takao Isogai
;  APPLICANT:  Tetsuo Nishikawa
;  APPLICANT:  Koji Hayashi
;  APPLICANT:  Kaoru Otsuka
;  APPLICANT:  Jun-Ichi Yamamoto
;  APPLICANT:  Shizuko Ishii
;  APPLICANT:  Tomoyasu Sugiyama
;  APPLICANT:  Ai Wakamatsu
;  APPLICANT:  Keiichi Nagai
;  APPLICANT:  Tetsuji Otsuki
;  APPLICANT:  Shin-Ichi Funahashi
;  APPLICANT:  Chiaki Senoo
;  APPLICANT:  Jun-Ichi Nezu
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
;  FILE REFERENCE: 06501-099002
;  CURRENT APPLICATION NUMBER: US/10/060,065
;  CURRENT FILING DATE:  2002-01-29
;  PRIOR APPLICATION NUMBER: PCT/JP00/05061
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: US 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: US 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: JP 11-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-060-065-13

  Query Match            50.1%;  Score 1057.5;  DB 14;  Length 344;
  Best Local Similarity  63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 29
US-10-059-585-34
; Sequence 34, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
;  APPLICANT: Ota, Toshio
;  APPLICANT:  Isogai, Takao
;  APPLICANT:  Nishikawa, Tetsuo
;  APPLICANT:  Hayashi, Koji
;  APPLICANT:  Otsuka, Kaoru
;  APPLICANT:  Yamamoto, Jun-ichi
;  APPLICANT:  Ishii, Shizuko
;  APPLICANT:  Sugiyama, Tomoyasu
;  APPLICANT:  Wakamatsu, Ai
;  APPLICANT:  Nagai, Keiichi
;  APPLICANT:  Otsuki, Tetsuji
;  APPLICANT:  Funahashi, Shin-Ichi
;  APPLICANT:  Senoo, Chiaki
;  APPLICANT:  Nezu, Jun-Ichi
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
;  TITLE OF INVENTION:  KINASE/PROTEIN PHOSPHATASE
;  FILE REFERENCE: 06501-098001
;  CURRENT APPLICATION NUMBER: US/10/059,585
```

```
;  CURRENT FILING DATE:  2002-01-29
;  PRIOR APPLICATION NUMBER: PCT/JP00/05060
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: US 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: US 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 11-248036
;  PRIOR FILING DATE: 1999-07-29
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-059-585-34

  Query Match             50.1%;  Score 1057.5;  DB 14;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:          | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:|||:||:|||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 30
US-10-171-311-214
; Sequence 214, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Chen, Yan
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Monahan, John
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Hoersh, Sebastian
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;  TITLE OF INVENTION:  OF CERVICAL CANCER
;  FILE REFERENCE: MRI-035
;  CURRENT APPLICATION NUMBER: US/10/171,311
;  CURRENT FILING DATE:  2002-06-12
;  PRIOR APPLICATION NUMBER: US 60/298,159
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,155
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/335,936
;  PRIOR FILING DATE: 2001-11-14
;  NUMBER OF SEQ ID NOS: 238
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 214
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-171-311-214

  Query Match             50.1%;  Score 1057.5;  DB 14;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:          | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:|||:||:|||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 31
US-10-295-027-203
; Sequence 203, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
```

```
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 203
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-295-027-203

  Query Match             50.1%;  Score 1057.5;  DB 15;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 32
US-10-173-999-101
; Sequence 101, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David H.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
;  TITLE OF INVENTION:  and Methods of Screening for Modulators of Ovarian
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 018501-002420US
;  CURRENT APPLICATION NUMBER: US/10/173,999
;  CURRENT FILING DATE:  2002-06-17
;  PRIOR APPLICATION NUMBER: US 60/299,234
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: US 60/315,287
;  PRIOR FILING DATE: 2001-08-27
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/372,246
;  PRIOR FILING DATE: 2001-04-12
;  NUMBER OF SEQ ID NOS: 163
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-173-999-101

  Query Match             50.1%;  Score 1057.5;  DB 15;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 33
US-10-188-832-93
; Sequence 93, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David H.
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
;  TITLE OF INVENTION:  and Methods of Screening for Modulators of Bladder
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 018501-002330US
;  CURRENT APPLICATION NUMBER: US/10/188,832
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: US 60/302,814
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: US 60/310,099
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: US 60/343,705
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/372,246
;  PRIOR FILING DATE: 2002-04-12
;  NUMBER OF SEQ ID NOS: 207
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-188-832-93

  Query Match             50.1%;  Score 1057.5;  DB 15;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
```

```
Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 34
US-10-734-126-3
; Sequence 3, Application US/10734126
; Publication No. US20040265852A1
;  GENERAL INFORMATION:
;       APPLICANT: Plowman, Gregory
;                  Mossie, Kevin
;       TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
;                           AND/OR AUR-2 RELATED DISORDERS
;       NUMBER OF SEQUENCES: 39
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street
;                    Suite 4700
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: U.S.A.
;            ZIP: 90071-2066
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;              ;          storage
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: IBM P.C. DOS 5.0
;            SOFTWARE: FastSEQ for Windows 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/734,126
;            FILING DATE: 15-Dec-2003
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/784,332
;            FILING DATE: 16-Feb-2001
;            APPLICATION NUMBER: 09/283,011
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 09/012,135
;            FILING DATE: January 22, 1998
;            APPLICATION NUMBER: 08/755,728
;            FILING DATE: November 25, 1996
;            APPLICATION NUMBER: 60/023,943
;            FILING DATE: August 14, 1996
;            APPLICATION NUMBER: 60/008,809
;            FILING DATE: December 18, 1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Warburg, Richard J.
;            REGISTRATION NUMBER: 32,327
;            REFERENCE/DOCKET NUMBER: 231/282
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (213) 489-1600
;            TELEFAX: (213) 955-0440
;            TELEX: 67-3510
;  INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 344 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-734-126-3

  Query Match             50.1%;  Score 1057.5;  DB 16;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 9.9e-70;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| ||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 35
US-10-429-849-4
; Sequence 4, Application US/10429849
; Publication No. US20040029157A1
; GENERAL INFORMATION:
;  APPLICANT: TATSUKA, MASAAKI
;  APPLICANT:  TERADA, YASUHIKO
;  TITLE OF INVENTION: CELL CYCLE-REGULATING PROTEINS
;  FILE REFERENCE: 050499/0102
;  CURRENT APPLICATION NUMBER: US/10/429,849
;  CURRENT FILING DATE:  2003-05-06
;  PRIOR APPLICATION NUMBER: US/09/485,534
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: PCT/JP98/03641
;  PRIOR FILING DATE: 1998-08-17
;  PRIOR APPLICATION NUMBER: 235371/1997
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Rattus sp.
US-10-429-849-4

  Query Match             49.8%;  Score 1050;  DB 15;  Length 343;
  Best Local Similarity   57.6%;  Pred. No. 3.5e-69;
  Matches  204;  Conservative   51;  Mismatches   73;  Indels   26;  Gaps    4;

Qy         36 PLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNT 95
              | |  |  :|  |  :   ||| |: :  |:  : : |:                 :|:
Db          9 PWPYGSKTSQSGL--NTLPQRV-LRKEPAVTPAQALMNR-----------------SNS 47

Qy         96 QKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQS 155
              | :  |     ||      |   |       :: : :::|||||||||||||||||||:|
Db         48 QSTAVPGQKLTENKGATALQGSQ-----SRQPFTIDNFEIGRPLGKGKFGNVYLAREKKS 102

Qy        156 KFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPL 215
```

```
              :||:|||:|||:|:|| |||||||||:|||:||:|||||:|| ||:|  |:||||||||
Db        103 RFIVALKILFKSQIEKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPR 162

Qy        216 GTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADF 275
              | :|:|||   |||||||| : ||:|| ||| |:|||||||||||||  ||||||||
Db        163 GELYKELQKSGTFDEQRTATIMEELSDALMYCHKKKVIHRDIKPENLLLGLQGELKIADF 222

Qy        276 GWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQ 335
              |||||||| || |:|||||||||||||||:| |||| :|||||| :|| ||||: ::
Db        223 GWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHS 282

Qy        336 ETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
              |||:|| :|:  ||  :  ||:||||:||||||||  | :|  |||: ||| :
Db        283 ETYRRIVKVDLKFPSSMPLGAKDLISKLLKHNPSQRLPLEQVSAHPWVRANSRR 336


RESULT 36
US-09-974-298-136
; Sequence 136, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Huei-Mei
;  TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
;  FILE REFERENCE: PA-0037 P
;  CURRENT APPLICATION NUMBER: US/09/974,298
;  CURRENT FILING DATE:  2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,331
;  PRIOR FILING DATE: 2000-05-10
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 136
;   LENGTH: 347
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020156263A1 002940CD1
US-09-974-298-136

  Query Match             49.6%;  Score 1046;  DB 9;  Length 347;
  Best Local Similarity   62.7%;  Pred. No. 7e-69;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   26;  Gaps    6;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |          |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTAT---YITELANALSY 246
              |||||||| ||:|  |:|||||||| | :|:||||   ||||||||    : |||:|| |
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATVRAIMEELADALMY 193

Qy        247 CHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMH 306
              || |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||||
Db        194 CHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMH 253

Qy        307 DEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKH 366
              :|||||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||:|
Db        254 NEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRH 313

Qy        367 NPSQRPMLREVLEHPWITANSSK---PSNCQN 395
              |||:|  | :|  |||: ||| :   ||  |:
Db        314 NPSERLPLAQVSAHPWVRANSRRVLPPSALQS 345


RESULT 37
US-10-403-571-98
; Sequence 98, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
;  APPLICANT: Hopkins, Nancy
;  APPLICANT:  Golling, Gregory
;  APPLICANT:  Amsterdam, Adam
;  APPLICANT:  Sun, Zhoaxia
;  TITLE OF INVENTION: Developmental Mutations in Zebrafish
;  FILE REFERENCE: 01997/539002
;  CURRENT APPLICATION NUMBER: US/10/403,571
;  CURRENT FILING DATE:  2003-03-25
;  PRIOR APPLICATION NUMBER: US 60/368,760
;  PRIOR FILING DATE: 2002-03-29
;  NUMBER OF SEQ ID NOS: 159
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 98
;   LENGTH: 320
;   TYPE: PRT
;   ORGANISM: Danio rerio
US-10-403-571-98

  Query Match             49.5%;  Score 1045;  DB 15;  Length 320;
  Best Local Similarity   64.1%;  Pred. No. 7.6e-69;
  Matches  200;  Conservative   40;  Mismatches   62;  Indels   10;  Gaps    3;

Qy         94 NTQKSKQPLPSA------PENNPEEELAS--KQKNEESKKRQWALEDFEIGRPLGKGKFG 145
              | :   |  |||       | ||:    |   :   :|  :  :::||:|||||||||||
Db          6 NREPRVQQTPSAGVGPLRVEMNPDTHAVSGPGRVPVKSNSKVLSIDDFDIGRPLGKGKFG 65

Qy        146 NVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATR 205
              |||||||:: | ::||||||:|: | |||||||||:|||||||||||| | |||| ||
Db         66 NVYLARERKLKVVIALKVLFKSQMVKEGVEHQLRREIEIQSHLRHPNILRFYNYFHDDTR 125

Qy        206 VYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLG 265
              |:|||||| | :|:|||:   ||:||||||: |:::|| ||| |:|||||||||||||
Db        126 VFLILEYAPRGEMYKELQRYGHFDDQRTATYMEEVSDALQYCHEKKVIHRDIKPENLLLG 185

Qy        266 SAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVG 325
              |||||||||||||||| || |:|||||||||||||  |||||||||:|||||| |||
Db        186 YRGELKIADFGWSVHAPSLRRRTMCGTLDYLPPEMIEGHSHDEKVDLWSIGVLCYECLVG 245

Qy        326 KPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITA 385
              ||||   : ||||||::|:  ||  |:|||||||:||:|:|| |  || |:||  : |
Db        246 NPPFETRQHAETYKRITKVDLQFPKLVSEGARDLISKLLRHSPSMRLPLRSVMEHRGVKA 305

Qy        386 NSSK--PSNCQN 395
              || :  |  | :
Db        306 NSRRVLPPVCSS 317


Search completed: June 20, 2005, 20:06:52
Job time : 82.1867 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:57:12 ; Search time 138.649 Seconds
(without alignments)
1488.417 Million cell updates/sec

Title: US-10-734-126-4
Perfect score: 2110
Sequence: 1 MDRSKENCISGPVKATAPVG.....TANSSKPSNCQNKESASKQS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 21

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2110	100.0	403	1	STK6_HUMAN	O14965 homo sapien
2	1686	79.9	397	1	STK6_RAT	P59241 rattus norv
3	1661.5	78.7	386	2	Q8C3H8	Q8c3h8 mus musculu
4	1661	78.7	395	1	STK6_MOUSE	P97477 mus musculu
5	1652	78.3	395	2	Q7TNK2	Q7tnk2 mus musculu
6	1650	78.2	395	2	Q8BP87	Q8bp87 mus musculu
7	1320.5	62.6	408	1	ST6L_XENLA	Q91819 xenopus lae
8	1318.5	62.5	408	2	Q6DJK0	Q6djk0 xenopus lae
9	1295.5	61.4	407	1	STK6_XENLA	Q91820 xenopus lae
10	1204.5	57.1	405	2	Q6DBZ4	Q6dbz4 brachydanio
11	1071.5	50.8	361	2	Q9DF70	Q9df70 xenopus lae
12	1071	50.8	320	2	Q6NW76	Q6nw76 brachydanio
13	1069	50.7	368	2	Q6GPL3	Q6gpl3 xenopus lae
14	1067.5	50.6	361	2	Q8JG74	Q8jg74 xenopus lae
15	1067.5	50.6	361	2	Q6DE08	Q6de08 xenopus lae
16	1067.5	50.6	371	2	Q7ZYT9	Q7zyt9 xenopus lae
17	1061	50.3	345	1	AURB_MOUSE	O70126 mus musculu
18	1057.5	50.1	344	1	AURB_HUMAN	Q96gd4 homo sapien
19	1057	50.1	345	2	Q8C6C1	Q8c6c1 mus musculu
20	1050	49.8	343	1	AURB_RAT	O55099 rattus norv
21	1045	49.5	344	2	Q7YRC6	Q7yrc6 bos taurus

ALIGNMENTS

RESULT 1
STK6_HUMAN
ID STK6_HUMAN STANDARD; PRT; 403 AA.
AC O14965; O60445; O75873; Q9BQD6; Q9UPG5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Serine/threonine-protein kinase 6 (EC 2.7.1.37) (Serine/threonine
DE kinase 15) (Aurora/IPL1-related kinase 1) (Aurora-related kinase 1)
DE (hARK1) (Aurora-A) (Breast-tumor-amplified kinase).
GN Name=STK6; Synonyms=AIK, ARK1, AURA, BTAK, STK15;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97298083; PubMed=9153231; DOI=10.1074/jbc.272.21.13766;
RA Kimura M., Kotani S., Hattori T., Sumi N., Yoshioka T., Todokoro K.,
RA Okano Y.;
RT "Cell cycle-dependent expression and spindle pole localization of a
RT novel human protein kinase, Aik, related to Aurora of Drosophila and
RT yeast Ipl1.";
RL J. Biol. Chem. 272:13766-13771(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98183439; PubMed=9514916; DOI=10.1006/bbrc.1998.8250;
RA Shindo M., Nakano H., Kuroyanagi H., Shirasawa T., Mihara M.,
RA Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;
RT "cDNA cloning, expression, subcellular localization, and chromosomal
RT assignment of mammalian aurora homologues, aurora-related kinase (ARK)
RT 1 and 2.";
RL Biochem. Biophys. Res. Commun. 244:285-292(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast;
RX MEDLINE=98442657; PubMed=9771714; DOI=10.1038/2496;
RA Zhou H., Kuang J., Zhong L., Kuo W.-L., Gray J.W., Sahin A.,
RA Brinkley B.R., Sen S.;
RT "Tumour amplified kinase STK15/BTAK induces centrosome amplification,
RT aneuploidy and transformation.";
RL Nat. Genet. 20:189-193(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA Wang L., Thibodeau S.N.;
RT "Mutational analysis of the STK15 gene in human tumors.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052; DOI=10.1038/414865a;
RA Deloukas P., Matthews L.H., Ashurst J.L., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).

```
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cervix, Colon, Kidney, and Muscle;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   CELL-CYCLE REGULATION.
RX   MEDLINE=21895866; PubMed=11790771; DOI=10.1074/jbc.M108252200;
RA   Tanaka M., Ueda A., Kanamori H., Ideguchi H., Yang J., Kitajima S.,
RA   Ishigatsubo Y.;
RT   "Cell-cycle-dependent regulation of human aurora A transcription is
RT   mediated by periodic repression of E4TF1.";
RL   J. Biol. Chem. 277:10719-10726(2002).
RN   [8]
RP   REVIEW.
RX   MEDLINE=21306577; PubMed=11413462; DOI=10.1038/35048096;
RA   Nigg E.A.;
RT   "Mitotic kinases as regulators of cell division and its checkpoints.";
RL   Nat. Rev. Mol. Cell Biol. 2:21-32(2001).
CC   -!- FUNCTION: May play a role in cell cycle regulation during anaphase
CC       and/or telophase, in relation to the function of the
CC       centrosome/spindle pole region during chromosome segregation.
CC       Maybe involved in microtubule formation and/or stabilization. May
CC       play a key role during tumor development and progression.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes on centrosomes in interphase cells
CC       and at each spindle pole in mitosis.
CC   -!- TISSUE SPECIFICITY: Highly expressed in testis and weakly in
CC       skeletal muscle, thymus and spleen. Also highly expressed in
CC       colon, ovarian, prostate, neuroblastoma, breast and cervical
CC       cancer cell lines. Expression is cell-cycle regulated, low in
CC       G1/S, accumulates during G2/M, and decreases rapidly after.
CC   -!- PTM: Phosphorylated.
CC   -!- DISEASE: Defects in STK6 are responsible for numerical centrosome
CC       aberrations including aneuploidy.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in positions 105, 125, 129, 235 and 241.
CC   -!- CAUTION: Although authors have considered STK6 and STK15 as two
CC       different proteins, it is clear that they are the same protein.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D84212; BAA23592.1; ALT_FRAME.
DR   EMBL; AF008551; AAC12708.1; -.
DR   EMBL; AF011467; AAC23448.1; -.
DR   EMBL; AF011468; AAC63902.1; -.
DR   EMBL; AF195947; AAF29508.1; -.
DR   EMBL; AF195942; AAF29508.1; JOINED.
DR   EMBL; AF195943; AAF29508.1; JOINED.
DR   EMBL; AF195944; AAF29508.1; JOINED.
DR   EMBL; AF195945; AAF29508.1; JOINED.
DR   EMBL; AF195946; AAF29508.1; JOINED.
DR   EMBL; AL121914; CAC12717.1; -.
DR   EMBL; BC001280; AAH01280.1; -.
DR   EMBL; BC002499; AAH02499.1; -.
DR   EMBL; BC006423; AAH06423.1; -.
DR   EMBL; BC027464; AAH27464.1; -.
DR   PDB; 1MQ4; X-ray; A=120-391.
DR   PDB; 1MUO; X-ray; A=107-403.
DR   PDB; 1OL6; X-ray; A=122-403.
DR   PDB; 1OL7; X-ray; A=122-403.
DR   Genew; HGNC:11393; STK15.
DR   Genew; HGNC:11409; STK6.
DR   H-InvDB; HIX0015930; -.
DR   MIM; 602687; -.
DR   MIM; 603072; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005819; C:spindle; TAS.
DR   GO; GO:0004672; F:protein kinase activity; TAS.
DR   GO; GO:0007067; P:mitosis; TAS.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   3D-structure; ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      133    383       Protein kinase.
FT   NP_BIND     139    147       ATP (By similarity).
FT   BINDING     162    162       ATP (By similarity).
FT   ACT_SITE    256    256       Proton acceptor (By similarity).
FT   CONFLICT     31     31       F -> I (in Ref. 3).
FT   CONFLICT     57     57       V -> I (in Ref. 2).
SQ   SEQUENCE   403 AA;  45809 MW;  125F3594834CD157 CRC64;

  Query Match           100.0%;  Score 2110;  DB 1;  Length 403;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-126;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
```

```
Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              ||||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 2
STK6_RAT
ID STK6_RAT STANDARD; PRT; 397 AA.
AC P59241;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase 6 (EC 2.7.1.37) (Aurora-A) (ratAurA).
GN Name=Stk6; Synonyms=Stk15;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar Furth; TISSUE=Mammary gland;
RX MEDLINE=22119444; PubMed=12124350;
RA Goepfert T.M., Adigun Y.E., Zhong L., Gay J., Medina D.,
RA Brinkley W.R.;
RT "Centrosome amplification and overexpression of aurora A are early
RT events in rat mammary carcinogenesis.";
RL Cancer Res. 62:4115-4122(2002).
RN [2]
RP REVIEW.
RX MEDLINE=21306577; PubMed=11413462; DOI=10.1038/35048096;
RA Nigg E.A.;
RT "Mitotic kinases as regulators of cell division and its checkpoints.";
RL Nat. Rev. Mol. Cell Biol. 2:21-32(2001).
CC -!- FUNCTION: May play a role in cell cycle regulation during anaphase
CC   and/or telophase, in relation to the function of the
CC   centrosome/spindle pole region during chromosome segregation.
CC   Maybe involved in microtubule formation and/or stabilization. May
CC   play a key role during tumor development and progression.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Localizes on centrosomes in interphase cells
CC   and at each spindle pole in mitosis (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in mammary gland and tumor.
CC -!- INDUCTION: Activated by progesterone.
CC -!- PTM: Phosphorylated (By similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC   subfamily.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF537333; AAN06823.1; -.
DR HSSP; P31751; 1GZK.
DR RGD; 628895; Stk6.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Cell cycle; Phosphorylation;
KW Serine/threonine-protein kinase; Transferase.
FT DOMAIN 126 376 Protein kinase.
FT NP_BIND 132 140 ATP (By similarity).
FT BINDING 155 155 ATP (By similarity).
FT ACT_SITE 249 249 Proton acceptor (By similarity).
SQ SEQUENCE 397 AA; 44874 MW; 95DECA2198DCED85 CRC64;

```
  Query Match            79.9%;  Score 1686;  DB 1;  Length 397;
  Best Local Similarity  83.4%;  Pred. No. 1.8e-99;
  Matches  337;  Conservative  15;  Mismatches  44;  Indels  8;  Gaps  5;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||| |||| :| |||:| |  |||||||| :| | |:|   :|||||||||||| ||||| |
Db          1 MDRCKENCVSRPVKSTVPF-GPKRVLVTEQIPSQHPGSASSGQAQRVLCPSN-SQRVPPQ 58

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||| |:  |||     ||| | | | || |:| |||:|| |:|  || |:|  | ||
Db         59 AQKPVAGQKPV----LKQLPAASGPRPASR-LSNPQKSEQPQPAASGNNSEKEQTSIQKT 113

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |:|||| |||||||||||||||||||||||||||||||||||| ||||||||||||
Db        114 EDSKKRQWTLEDFDIGRPLGKGKFGNVYLAREKQSKFILALKVLFKVQLEKAGVEHQLRR 173

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 233

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||| ||||
Db        234 ANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTLCGTLDYQPPEM 293

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              |||||||||||||||||||||||| |||||:||||||:||||||||||||||||||||||
Db        294 IEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDLI 353

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQN-KESASKQS 403
              ||||||| |||  | ||||||||| ||||||    | ||: || |
Db        354 SRLLKHNSSQRLTLAEVLEHPWIKANSSKPPTGHNSKEATSKSS 397

RESULT 3
Q8C3H8
ID  Q8C3H8  PRELIMINARY;  PRT;  386 AA.
AC  Q8C3H8;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched
DE  library, clone:D830019E15 product:serine/threonine kinase 6, full
DE  insert sequence.
GN  Name=Stk6;
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Heart;
RX  MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA  Carninci P., Hayashizaki Y.;
RT  "High-efficiency full-length cDNA cloning.";
RL  Meth. Enzymol. 303:19-44(1999).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Heart;
RX  MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA  RIKEN FANTOM Consortium;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Heart;
RA  The FANTOM Consortium,
RA  the RIKEN Genome Exploration Research Group Phase I & II Team;
RT  "Analysis of the mouse transcriptome based on functional annotation of
RT  60,770 full-length cDNAs.";
RL  Nature 420:563-573(2002).
```

```
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AK085861; BAC39557.1; -.
DR   HSSP; O14965; 1OL6.
DR   MGD; MGI:894678; Stk6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   386 AA;  43968 MW;  31970E9CDE31C6A7 CRC64;

  Query Match             78.7%;  Score 1661.5;  DB 2;  Length 386;
  Best Local Similarity   84.2%;  Pred. No. 6.2e-98;
  Matches  330;  Conservative   16;  Mismatches   37;  Indels    9;  Gaps    5;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||| |||| :| ||| | |  |||||||| :| | ||    :|||||||||||| ||||| |
Db          1 MDRCKENCVSRPVKTTVPF-GPKRVLVTEQIPSQNLGSASSGQAQRVLCPSN-SQRVPSQ 58

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||| :  ||       ||| | ||| |||| ||| ||::|  |:|  |: |:| || ||
Db         59 AQKLGAGQKPA----PKQLPAASVPRPVSR-LNNPQKNEQ--PAASGNDSEKEQASLQKT 111

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |::||||| |||||:|||||||||||||||||:|||||||||||||| ||||| |||||||

Db        112 EDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLRR 171

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 231

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              |||||||||||||||||||||||||||| ||||||||||||||||||:||||||||||||
Db        232 ANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPEM 291

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              |||||||||||||||||||||||| ||||:|||||:|||||||||||||||||||||||||
Db        292 IEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDLI 351

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSN 392
              ||||||| |||  | |||||||| ||||||||
Db        352 SRLLKHNASQRLTLAEVLEHPWIKANSSKPSN 383


RESULT 4
STK6_MOUSE
ID   STK6_MOUSE     STANDARD;      PRT;   395 AA.
AC   P97477; O35624; Q91YU4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase 6 (EC 2.7.1.37) (Aurora-family kinase
DE   1) (Aurora/IPL1-related kinase 1) (Ipl1- and aurora-related kinase 1)
DE   (Aurora-A) (Serine/threonine kinase Ayk1).
GN   Name=Stk6; Synonyms=Ark1, Ayk1, Iak1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=BALB/c; TISSUE=Testis;
RX   MEDLINE=97392693; PubMed=9245792; DOI=10.1083/jcb.138.3.643;
RA   Gopalan G., Chan C.S.M., Donovan P.J.;
RT   "A novel mammalian, mitotic spindle-associated kinase is related to
RT   yeast and fly chromosome segregation regulators.";
RL   J. Cell Biol. 138:643-656(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Testis;
RX   MEDLINE=97349097; PubMed=9205101; DOI=10.1038/sj.onc.1201144;
RA   Yanai A., Arama E., Kilfin G., Motro B.;
RT   "ayk1, a novel mammalian gene related to Drosophila aurora centrosome
RT   separation kinase, is specifically expressed during meiosis.";
RL   Oncogene 14:2943-2950(1997).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=BALB/c;
RX   MEDLINE=98183439; PubMed=9514916; DOI=10.1006/bbrc.1998.8250;
RA   Shindo M., Nakano H., Kuroyanagi H., Shirasawa T., Mihara M.,
RA   Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;
RT   "cDNA cloning, expression, subcellular localization, and chromosomal
RT   assignment of mammalian aurora homologues, aurora-related kinase (ARK)
RT   1 and 2.";
RL   Biochem. Biophys. Res. Commun. 244:285-292(1998).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

```
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   REVIEW.
RX   MEDLINE=21306577; PubMed=11413462; DOI=10.1038/35048096;
RA   Nigg E.A.;
RT   "Mitotic kinases as regulators of cell division and its checkpoints.";
RL   Nat. Rev. Mol. Cell Biol. 2:21-32(2001).
CC   -!- FUNCTION: May play a role in cell cycle regulation during anaphase
CC       and/or telophase, in relation to the function of the
CC       centrosome/spindle pole region during chromosome segregation.
CC       Maybe involved in microtubule formation and/or stabilization.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes on centrosomes in interphase cells
CC       and at each spindle pole in mitosis (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P97477-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P97477-2; Sequence=VSP_004871;
CC         Note=May be less abundant or less stable;
CC   -!- TISSUE SPECIFICITY: Highly expressed in testis, in about one third
CC       of the seminiferous tubules. Expression is restricted to specific
CC       spermatocytes nearing completion of prophase, with levels falling
CC       off on transition to elongated spermatids. Highly expressed in the
CC       ovary, expression in the oocyte starts around the transition to
CC       large growing follicle. Abundant expression is seen in the
CC       proliferating granulosa and thecal cells of the growing follicle,
CC       and in the young corpus luteum. Very weakly expressed in spleen
CC       and intestine.
CC   -!- DEVELOPMENTAL STAGE: At 7.5-9.5 dpc expressed evenly all over the
CC       embryo. At later stages, expression is mainly restricted to
CC       proliferating zones. The highest levels of expression at mid-
CC       embryonic development (13.5 dpc) were observed in the liver, lung,
CC       kidney and back (trapezius) muscle and all regions in active
CC       proliferation.
CC   -!- PTM: Phosphorylated (By similarity).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U80932; AAB62982.1; -.
DR   EMBL; AF007817; AAB63205.1; -.
DR   EMBL; U69106; AAC12682.1; -.
DR   EMBL; BC005425; AAH05425.1; -.
DR   EMBL; BC014711; AAH14711.1; -.
DR   HSSP; P31751; 1GZK.
DR   MGD; MGI:894678; Stk6.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Alternative splicing; ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      124    374       Protein kinase.
FT   NP_BIND     130    138       ATP (By similarity).
FT   BINDING     153    153       ATP (By similarity).
FT   ACT_SITE    247    247       Proton acceptor (By similarity).
FT   VARSPLIC      1      1       M -> MAVEGEPGCCKRIGKAVWRRGDM (in isoform
FT                                2).
FT                                /FTId=VSP_004871.
FT   CONFLICT    234    234       A -> T (in Ref. 1).
SQ   SEQUENCE   395 AA;  44772 MW;  26B6B65105A1A812 CRC64;

  Query Match             78.7%;  Score 1661;  DB 1;  Length 395;
  Best Local Similarity   82.4%;  Pred. No. 6.9e-98;
  Matches  333;  Conservative   18;  Mismatches   43;  Indels   10;  Gaps    6;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||| |||| :| ||| | |  |||||||| : | ||    :|||||||||||| ||||| |
Db          1 MDRCKENCVSRPVKTTVPF-GPKRVLVTEQIPSQNLGSASSGQAQRVLCPSN-SQRVPSQ 58

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||| :  ||      ||| | ||| |||| ||| ||::|  |:|  |: |:| || ||
Db         59 AQKLGAGQKPA----PKQLPAASVPRPVSR-LNNPQKNEQ--PAASGNDSEKEQASLQKT 111

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |::||||| ||||:||||||||||||||:||||||||||| ||||| |||||||
Db        112 EDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLRR 171

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 231

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              |||||||||||||||||||||||| ||||||||||||||||||:||||||||||
Db        232 ANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPEM 291

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              |||||||||||||||||||||||| |||||:||||||:||||||||||||||||||||
Db        292 IEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDLI 351

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSK-PSNCQNKESASKQS 403
              ||||||| |||  | |||||||| ||||| |:   :||  || |
Db        352 SRLLKHNASQRLTLAEVLEHPWIKANSSKPPTGHTSKEPTSKSS 395


RESULT 5
Q7TNK2
ID   Q7TNK2      PRELIMINARY;      PRT;   395 AA.
AC   Q7TNK2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Serine/threonine protein kinase 6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   Zhou G., Li W., Yu L.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY336976; AAQ16152.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
```

```
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   395 AA;  44771 MW;  271DC321221A8DD9 CRC64;

  Query Match           78.3%;  Score 1652;  DB 2;  Length 395;
  Best Local Similarity   82.2%;  Pred. No. 2.6e-97;
  Matches  332;  Conservative   18;  Mismatches   44;  Indels   10;  Gaps    6;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||| || |:| ||| | |  |||||||:| | ||    :||||||||||| ||||| |
Db          1 MDRCKEICVSRPVKTTVPF-GPKRVLVTEQIPSQNLGSASSGQAQRVLCPSN-SQRVPSQ 58

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||| :  ||      ||| | ||| |||| ||| ||::|  |:|  |: |:| || ||
Db         59 AQKLGAGQKPA----PKQLPAASVPRPVSR-LNNPQKNEQ--PAASGNDSEKEQASLQKT 111

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |::||||| ||||:||||||||||||||:|||||||||||| ||||| |||||||
Db        112 EDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLRR 171

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 231

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              |||||||||||||||||||||||| ||||||||||||||||||:||||||||||||
Db        232 ANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPEM 291

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPPEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              |||||||||||||||||||||||| |||||:||||||:||||||||||||||||||||||
Db        292 IEGRMHDEKVDLWSLGVLCYEFLVGMPPPFEAHTYQETYRRISRVEFTFPDFVTEGARDLI 351

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSK-PSNCQNKESASKQS 403
              ||||||| |||  | |||||||| ||||| |:   :||  || |
Db        352 SRLLKHNASQRLTLAEVLEHPWIKANSSKPPTGHTSKEPTSKSS 395


RESULT 6
Q8BP87
ID   Q8BP87      PRELIMINARY;      PRT;   395 AA.
AC   Q8BP87;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched
DE   library, clone:5730433I06 product:serine/threonine kinase 6, full
DE   insert sequence.
GN   Name=Stk6;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AK077513; BAC36838.1; -.
DR   HSSP; O14965; 1OL6.
DR   MGD; MGI:894678; Stk6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   395 AA;  44813 MW;  3B081DFE3DC52667 CRC64;

  Query Match           78.2%;  Score 1650;  DB 2;  Length 395;
  Best Local Similarity   81.9%;  Pred. No. 3.4e-97;
  Matches  331;  Conservative   18;  Mismatches   45;  Indels   10;  Gaps    6;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||| ||||:| ||| | |  |||||||:| | ||    :||||||||||| ||||| |
```

```
Db          1 MDRCKENCVSRPVKTTVPF-GPKRVLVTEQIPSQNLGSASSGQAQRVLCPSN-SQRVPSQ 58

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||  :  ||       ||| | ||| |||| ||| ||::|  |:|  |: |:| || ||
Db         59 AQKLGAGQKPA----PKQLPAASVPRPVSR-LNNPQKNEQ--PAASGNDSEKEQASLQKT 111

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |::|||| |||:|||||||||||||||:|||||||||||| ||||| |||||||
Db        112 EDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLRR 171

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        172 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTPTYITEL 231

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              |||||||||||||||||||||||| |||||||||||||||||:||||||||||
Db        232 ANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPEM 291

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              |||||||||||||||||||| || |||||:||||||:||||||||||||||||||||||
Db        292 IEGRMHDEKVDLWSLGVLCYEFQVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDLI 351

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSK-PSNCQNKESASKQS 403
              ||||||| |||  | ||||||||| ||||| |:   :||  || |
Db        352 SRLLKHNASQRLTLAEVLEHPWIKANSSKPPTGHTSKEPTSKSS 395


RESULT 7
ST6L_XENLA
ID   ST6L_XENLA     STANDARD;      PRT;   408 AA.
AC   Q91819;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase Eg2-like (EC 2.7.1.37) (p46XlEg22).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Egg;
RA   Roghi C., Le Guellec R., Paris J., Couturier A., Philippe M.;
RT   "Eg2, selected by differential screening encodes a new Xenopus protein
RT   kinase family.";
RL   Submitted (OCT-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis
CC       especially from prophase through anaphase. Partially colocalised
CC       with gamma tubulin in the centrosome, from S to M phase (By
CC       similarity).
CC   -!- PTM: Phosphorylated (By similarity).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z17206; CAA78914.1; ALT_INIT.
DR   HSSP; P06244; 1FOT.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      140    390       Protein kinase.
FT   NP_BIND     146    154       ATP (By similarity).
FT   BINDING     169    169       ATP (By similarity).
FT   ACT_SITE    263    263       Proton acceptor (By similarity).
SQ   SEQUENCE   408 AA;  46461 MW;  97F6A69C7357AEE8 CRC64;

  Query Match            62.6%;  Score 1320.5;  DB 1;  Length 408;
  Best Local Similarity  66.5%;  Pred. No. 3.5e-76;
  Matches  266;  Conservative   47;  Mismatches   74;  Indels   13;  Gaps    6;

Qy          1 MDRS-KENCISGPVKATAPV-GGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVP 58
              |:|: |||      ||   |:  ||||: |:|    |   ||    |||:| |||  |||
Db          1 MERAVKENHKPSNVKVFHPMTEGPKRIPVSQPPSTQVRPPVTGVSAQRILGPSNVPQRVM 60

Qy         59 LQAQK-LVSSHKPVQNQKQKQLQATSVPHPVSRPLN-NTQKSKQPLPSAPENNPEEELAS 116
              :|||| ::|: ||     |  |:  :  |  |:|   |  :: |   :  | |:: ::
Db         61 MQAQKPVLSNQKPT---AQGLLRPATHGHQTSKPQGPNENRNPQQTSHSSTPNMEKKGST 117

Qy        117 KQ------KNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLE 170
              |       || ||:|| |||||||||||||||||||::|||||||||||:|||
Db        118 DQGKTLAVPKEEGKKKQWCLEDFEIGRPLGKGKFGNVYLARERESKFILALKVLFKSQLE 177

Qy        171 KAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDE 230
              |||||||||||||||||||||||||||||:|||||:||| | ::||||| ::||:
Db        178 KAGVEHQLRREVEIQSHLRHPNILRLYGYFHDASRVYLILDYAPGGELFRELQKCTRFDD 237

Qy        231 QRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLC 290
              ||:| || :|| || |||||:|||||||||||||| ||||||||||||||||||||
Db        238 QRSALYIKQLAEALLYCHSKKVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTLC 297

Qy        291 GTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPD 350
              |||||||||||||||||||||||||||||||||||||||||||| :|:||||:|||:||| :|
Db        298 GTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFETDTHQETYRRISKVEFQYPP 357

Qy        351 FVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSKP 390
              :|:| |:||:|:|||||: |  |: |||||||  || :|
Db        358 YVSEEAKDLVSKLLKHNPNHRLPLKGVLEHPWIVKNSQQP 397


RESULT 8
Q6DJK0
ID   Q6DJK0      PRELIMINARY;      PRT;   408 AA.
AC   Q6DJK0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   LOC398349 protein.
GN   Name=LOC398349;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
```

```
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Klein S., Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC075177; AAH75177.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   408 AA;  46477 MW;  FD4855CC6708E5E7 CRC64;

  Query Match             62.5%;  Score 1318.5;  DB 2;  Length 408;
  Best Local Similarity   66.5%;  Pred. No. 4.7e-76;
  Matches  266;  Conservative   46;  Mismatches   75;  Indels   13;  Gaps    6;

Qy          1 MDRS-KENCISGPVKATAPV-GGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVP 58
              |:|: |||     ||   |:  ||||: |:|     |   ||    |||:| |||  |||
Db          1 MERAVKENHKPSNVKVFHPMTEGPKRIPVSQPPSTQVRPPVTGVSAQRILGPSNVPQRVM 60

Qy         59 LQAQK-LVSSHKPVQNQKQKQLQATSVPHPVSRPLN-NTQKSKQPLPSAPENNPEEELAS 116
              :|||| ::|: ||     | |:  :  |  |:|   |  :: |    :   | |:: ::
Db         61 MQAQKPVLSNQKPT---AQGLLRPATHGHQTSKPQGPNENRNPQQTSHSSTPNMEKKGST 117

Qy        117 KQ------KNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLE 170
               |       || ||:|| ||||||||||||||||||||||::||||||||||||:|||
Db        118 DQGKTLAVPKEEGKKKQWCLEDFEIGRPLGKGKFGNVYLARERESKFILALKVLFKSQLE 177

Qy        171 KAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDE 230
              |||||||||||||||||||||||||||||||| |||||:||| | ::||||| ::||:
Db        178 KAGVEHQLRREVEIQSHLRHPNILRLYGYFHDACRVYLILDYAPGGELFRELQKCTRFDD 237

Qy        231 QRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLC 290
              ||:| || :|| || |||||:|||||||||||||| ||||||||||||||||||||||
Db        238 QRSALYIKQLAEALLYCHSKKVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTLC 297

Qy        291 GTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPD 350
              |||||||||||||||||||||||||||||||||||| :|:||||:|||:||| :|
Db        298 GTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPPETDTHQETYRRISKVEFQYPP 357

Qy        351 FVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSKP 390
              :|:| |:||:|:|||||: |  |: |||||||  || :|
Db        358 YVSEEAKDLVSKLLKHNPNHRLPLKGVLEHPWIVKNSQQP 397


RESULT 9
STK6_XENLA
ID   STK6_XENLA     STANDARD;      PRT;   407 AA.
AC   Q91820;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase Eg2 (EC 2.7.1.37) (pEg2) (p46Eg265).
GN   Name=EG2;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Egg;
RX   MEDLINE=98122855; PubMed=9454730;
RA   Roghi C., Giet R., Uzbekov R., Morin N., Chartrain I., Le Guellec R.,
RA   Couturier A., Doree M., Phillippe M., Prigent C.;
RT   "The Xenopus protein kinase pEg2 associates with the centrosome in a
RT   cell cycle-dependent manner, binds to the spindle microtubules and is
RT   involved in bipolar mitotic spindle assembly.";
RL   J. Cell Sci. 111:557-572(1998).
CC   -!- FUNCTION: Associates with the centrosome in a cell-cycle dependent
CC       manner and invades the microtubules at the poles of the spindle
CC       during mitosis suggesting that it may be involved in the corect
CC       formation of bipolar mitotic spindles.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis
CC       especially from prophase through anaphase. Partially colocalised
CC       with gamma tubulin in the centrosome, from S to M phase.
CC   -!- TISSUE SPECIFICITY: Highly expressed in ovary and testis.
CC   -!- PTM: Phosphorylated.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z17207; CAA78915.1; -.
DR   PIR; S52243; S52243.
DR   HSSP; Q16539; 1KV2.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      140    390       Protein kinase.
FT   NP_BIND     146    154       ATP (By similarity).
FT   BINDING     169    169       ATP (By similarity).
FT   ACT_SITE    263    263       Proton acceptor (By similarity).
SQ   SEQUENCE   407 AA;  46372 MW;  DE1628A2C6D11277 CRC64;

  Query Match             61.4%;  Score 1295.5;  DB 1;  Length 407;
  Best Local Similarity   64.9%;  Pred. No. 1.4e-74;
  Matches  264;  Conservative   38;  Mismatches   72;  Indels   33;  Gaps    6;
```

```
Qy          1 MDRS-KENCISGPVKATAPV-GGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVP 58
              |:|: |||      ||   |:  | ||: | |     |   | :  |||:| |||  |||
Db          1 MERAVKENHKPSNVKIFHPMTEGAKRIPVNQPQSTQFRPPGTAVSAQRILGPSNVPQRVL 60

Qy         59 LQAQK-LVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASK 117
               |||| ::|| ||           | :|    ||     :| :|       ||::   |
Db         61 AQAQKPILSSQKP----------TTQIP---LRPATQGHQSSKPQGPNENRNPQQTSHSS 107

Qy        118 QKN----------------EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILA 160
              |                  || ||:|| |||||||||||||||||||::||||||
Db        108 TPNVEKKGSTDQGKTSAVPKEEGKKKQWCLEDFEIGRPLGKGKFGNVYLARERESKFILA 167

Qy        161 LKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYR 220
              ||||||:|||||||||||||||||||||||||||||||||:||||||:||| | ::|
Db        168 LKVLFKSQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDASRVYLILDYAPGGELFR 227

Qy        221 ELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVH 280
              |||| ::||:||:| || :|| || |||||:||||||||||||| ||||||||||||
Db        228 ELQKCTRFDDQRSAMYIKQLAEALLYCHSKKVIHRDIKPENLLLGSNGELKIADFGWSVH 287

Qy        281 APSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKR 340
              |||||||||||||||||||||||||| |||||||||||||||||| :|:||||:|
Db        288 APSSRRTTLCGTLDYLPPEMIEGRMHDETVDLWSLGVLCYEFLVGKPPFETDTHQETYRR 347

Qy        341 ISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANS 387
              ||:||| :| :|:| ||||:|:||||||: |  |: |||||||  ||
Db        348 ISKVEFQYPPYVSEEARDLVSKLLKHNPNHRLPLKGVLEHPWIIKNS 394
```

RESULT 10
Q6DBZ4
ID Q6DBZ4 PRELIMINARY; PRT; 405 AA.
AC Q6DBZ4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Zgc:100912 protein.
GN Name=zgc:100912;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RA Director MGC Project;
RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.

CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; BC078304; AAH78304.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 405 AA; 45447 MW; B04793FC827BD489 CRC64;

Query Match 57.1%; Score 1204.5; DB 2; Length 405;
Best Local Similarity 60.7%; Pred. No. 8.6e-69;
Matches 242; Conservative 45; Mismatches 85; Indels 27; Gaps 6;

```
Qy          4 SKENCISGPVKATAPVGGPKRVLVTQ--QFPCQNPLPVNSGQAQ----------RVLCPS 51
              |::  |  |         ||||| |||  | |  ||      | ||          :  ||
Db         10 SRDLKIHRPENEKVSAAGPKRVPVTQSVQKPVSNPHTRVLGAAQGPQRVQRPVGKTSCPG 69

Qy         52 NSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNT-QKSKQPLPSAPENNP 110
              : :  | |  |         |||  : | :          |:|   | | : |||  |    ::
Db         70 DQNTR-PEQ-------HKPAAHSKPQP------KPPLSAETNKTAEPSKQDKPHQTPSST 115

Qy        111 EEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLE 170
                    ::   : :| |: | ||:|:||| |||||||:|||||||:|:|||||||||| |||
Db        116 TSTSSNTSGSSKSSKKAWTLENFDIGRALGKGKFGSVYLAREQQTKFILALKVLFKKQLE 175

Qy        171 KAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDE 230
              |||||||||||||||||||||||||||||||||||| |||||||:|| | :| |||:   ||:
Db        176 KAGVEHQLRREVEIQSHLRHPNILRLYGYFHDAARVYLILEFAPKGELYGELQRCGTFDD 235

Qy        231 QRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLC 290
              ||:|||| |||:|| ||||| ||||||||||||||: ||||||||||||| |||||:|||
Db        236 QRSATYIMELADALRYCHSKAVIHRDIKPENLLLGANGELKIADFGWSVHTPSSRRSTLC 295

Qy        291 GTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPD 350
              |||||||||||||: |||||||||||||||||||:|||| :::|||::|||||||:|
Db        296 GTLDYLPPEMIEGKTHDEKVDLWSLGVLCYEFLVGRPPFETKSHEETYRKISRVEFTYPA 355

Qy        351 FVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
              |: |:||||:|||||||  |  :: |:||||:  ||:|
Db        356 HVSNGSRDLINRLLKHNPMHRLPIQGVMEHPWVVENSTK 394
```

RESULT 11
Q9DF70
ID Q9DF70 PRELIMINARY; PRT; 361 AA.
AC Q9DF70;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Protein kinase AIRK2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20453769; PubMed=10996078; DOI=10.1016/S0960-9822(00)00673-4;
RA Adams R.R., Wheatleya S.P., Gouldsworthy A.M., Kandels-Lewis S.E.,
RA Carmena M., Smythe C., Gerloff D.L., Earnshaw W.C.;
RT "INCENP binds the Aurora-related kinase AIRK2 and is required to

```
RT   target it to chromosomes, the central spindle and cleavage furrow.";
RL   Curr. Biol. 10:1075-1078(2000).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AF292096; AAG10787.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   361 AA;  41777 MW;  BE577A9135FD0C93 CRC64;

  Query Match             50.8%;  Score 1071.5;  DB 2;  Length 361;
  Best Local Similarity   55.9%;  Pred. No. 2.2e-60;
  Matches  213;  Conservative   60;  Mismatches   63;  Indels   45;  Gaps    8;

Qy         35 NPLPVNSGQAQRVLCPSN--SSQRVPLQAQKLVSSH-KPVQNQKQKQLQ--------ATS 83
              ||    |    :   ||:  ::||| |: :  ||:   |  |   :: |        ::|
Db          8 NP----SSYTSKFTTPSSATAAQRV-LRKEPYVSTFTTPSDNLLAQRTQLSRITPSASSS 62

Qy         84 VPHPVS----RPLNNTQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPL 139
              ||  |:     |  ||  :: |                        |||: ::||:|||||
Db         63 VPGRVAVSTEMPSQNTALAEMP----------------------KRQFTIDDFDIGRPL 99

Qy        140 GKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGY 199
              |||||||||||||||:|||:|||||||:|||| ||||||||:|||||||||||||:| |
Db        100 GKGKFGNVYLAREKQNKFIMALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNY 159

Qy        200 FHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKP 259
              |||  |:||:||:|| | :|:|||| :|||||:||:: |||:|| ||| ::||||||||
Db        160 FHDRKRIYLMLEFAPRGELYKELQKHGRFDEQRSATFMEELADALHYCHERKVIHRDIKP 219

Qy        260 ENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLC 319
              ||||:|  ||||||||||||||| || |:|||||||||||||: ||||||||  ||||
Db        220 ENLLMGYKGELKIADFGWSVHAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLC 279

Qy        320 YEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLE 379
              |||||| |||:: :: ||::||  |:  || |:::|::||||:||:::| ||  |: |:|
Db        280 YEFLVGMPPFDSPSHTETHRRIVNVDLKFPPFLSDGSKDLISKLLRYHPPQRLPLKGVME 339

Qy        380 HPWITANSSK--PSNCQNKES 398
              |||: ||| :  |   |: :|
Db        340 HPWVKANSRRVLPPVYQSTQS 360


RESULT 12
Q6NW76
ID   Q6NW76      PRELIMINARY;      PRT;   320 AA.
AC   Q6NW76;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Serine/threonine kinase a.
GN   Name=stka;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC067695; AAH67695.1; -.
DR   HSSP; Q63450; 1A06.
DR   ZFIN; ZDB-GENE-020419-40; stka.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   320 AA;  36946 MW;  D1D9EC08556D1CF1 CRC64;

  Query Match             50.8%;  Score 1071;  DB 2;  Length 320;
  Best Local Similarity   64.5%;  Pred. No. 2e-60;
  Matches  202;  Conservative   43;  Mismatches   58;  Indels   10;  Gaps    3;

Qy         94 NTQKSKQPLPSA------PENNPEEELAS--KQKNEESKKRQWALEDFEIGRPLGKGKFG 145
              | :   |  |||       | ||:    |   :   :|  :  :::||:|||||||||||
Db          6 NREPRVQQTPSAGVGPLRVEMNPDTHAVSGPGRVPVKSNSKVLSIDDFDIGRPLGKGKFG 65

Qy        146 NVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATR 205
              |||||||:: | ::|||||||:|: | |||||||||:|||||||||||||| | |||| ||
Db         66 NVYLARERKLKVVIALKVLFKSQMVKEGVEHQLRREIEIQSHLRHPNILRFYNYFHDDTR 125

Qy        206 VYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLG 265
              |:|||||| | :|:|||:  :||:|||||: |:::|| ||| |:|||||||||||||||
Db        126 VFLILEYAPRGEMYKELQRYGRFDDQRTATYMEEVSDALQYCHEKKVIHRDIKPENLLLG 185

Qy        266 SAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVG 325
                |||||||||||||| || |:|||||||||||||  |||||||||:|||||| |||
Db        186 YRGELKIADFGWSVHAPSLRRRTMCGTLDYLPPEMIEGHSHDEKVDLWSIGVLCYECLVG 245

Qy        326 KPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITA 385
              ||||  :: ||||||::|:  ||  |:|||||||:||:|:|| |  || |:||||: |
```

```
Db        246 NPPPFETASHAETYKRITKVDLQFPKLVSEGARDLISKLLRHSPSMRLPLRSVMEHPWVKA 305

Qy        386 NSSK--PSNCQNK 396
              || :  |  | ::
Db        306 NSRRVLPPVCSSE 318


RESULT 13
Q6GPL3
ID   Q6GPL3      PRELIMINARY;      PRT;   368 AA.
AC   Q6GPL3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC83575 protein.
GN   Name=MGC83575;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC073103; AAH73103.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   368 AA;  42015 MW;  D1E79FB995701214 CRC64;

  Query Match           50.7%;  Score 1069;  DB 2;  Length 368;
  Best Local Similarity   59.1%;  Pred. No. 3.3e-60;
  Matches  207;  Conservative   56;  Mismatches   73;  Indels   14;  Gaps    5;

Qy         49 CPSNSSQRVPLQAQKLV--SSHKPVQNQKQKQLQATSVPHPVSRPL-NNTQKSKQPLPSA 105
              | |:||      : |    ||    |   :||   :    |    |   | : :  |:|
Db         12 CSSSSS-----SSSKFATPSSATAAQRVLRKQPYVSIFTTPSDNLLAQRAQLAPRVTPAA 66

Qy        106 PENNPEE-----ELASKQKN-EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFIL 159
                : |       ::||       |: ||:: ::||:|||||||||||||||||||:|||:
Db         67 SSSVPGRVALGTDVASHNTALAEAPKRKFTIDDFDIGRPLGKGKFGNVYLAREKQNKFIM 126

Qy        160 ALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVY 219
              |||||||:|||| |||||||||:|||||||||||||:| ||||  |:||:||:|| | :|
Db        127 ALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNYFHDRKRIYLMLEFAPRGELY 186

Qy        220 RELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSV 279
              :|||  :|||||:||:: |||:|| ||| ::|||||||||||||:|  ||||||||||
Db        187 KELQKHGRFDEQRSATFMEELADALQYCHERKVIHRDIKPENLLMGYKGELKIADFGWSV 246

Qy        280 HAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYK 339
              |||| || |:|||||||||||||: ||||||||  |||||||||| |||:: :: ||::
Db        247 HAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLCYEFLVGMPPFDSPSHSETHR 306

Qy        340 RISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
              ||  |:  || |::||::||||:||:::|:||  |: |:||||: ||| :
Db        307 RIVNVDLKFPPFLSEGSKDLISKLLRYHPAQRLPLKGVMEHPWVKANSRR 356


RESULT 14
Q8JG74
ID   Q8JG74      PRELIMINARY;      PRT;   361 AA.
AC   Q8JG74;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Aurora B.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22209728; PubMed=12221116;
RA   Bolton M.A., Lan W., Powers S.E., McCleland M.L., Kuang J.,
RA   Stukenberg P.T.;
RT   "Aurora B kinase exists in a complex with survivin and INCENP and its
RT   kinase activity is stimulated by survivin binding and
RT   phosphorylation.";
RL   Mol. Biol. Cell 13:3064-3077(2002).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY115554; AAM76715.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
```

```
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   361 AA;  41777 MW;  1FB77EF41E1D07E8 CRC64;

  Query Match           50.6%;  Score 1067.5;  DB 2;  Length 361;
  Best Local Similarity   55.6%;  Pred. No. 4e-60;
  Matches  212;  Conservative   61;  Mismatches   63;  Indels   45;  Gaps    8;

Qy         35 NPLPVNSGQAQRVLCPSN--SSQRVPLQAQKLVSSH-KPVQNQKQKQLQ--------ATS 83
              ||    |    :   ||:  ::||| |: :  ||:   |  |   :: |        ::|
Db          8 NP----SSYTSKFTTPSSATAAQRV-LRKEPYVSTFTTPSDNLLAQRTQLSRITPSASSS 62

Qy         84 VPHPVS----RPLNNTQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPL 139
              ||  |:     |  ||  :: |                        ||:: ::||:|||||
Db         63 VPGRVAVSTEMPSQNTALAEMP----------------------KRKFTIDDFDIGRPL 99

Qy        140 GKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGY 199
              |||||||||||||||:|||:|||||||:|||| |||||||||:|||||||||||||:| |
Db        100 GKGKFGNVYLAREKQNKFIMALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNY 159

Qy        200 FHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKP 259
              |||  |:||:||:|| | :|:|||| :|||||:||:: |||:|| ||| ::||||||||
Db        160 FHDRKRIYLMLEFAPRGELYKELQKHGRFDEQRSATFMEELADALHYCHERKVIHRDIKP 219

Qy        260 ENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLC 319
              ||||:|  |||||||||||||||| || |:|||||||||||||: ||||||||  ||||
Db        220 ENLLMGYKGELKIADFGWSVHAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLC 279

Qy        320 YEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLE 379
              |||||| |||:: :: ||::||  |:  || |:::|::||||:||:::| ||  |: |:|
Db        280 YEFLVGMPPFDSPSHTETHRRIVNVDLKFPPFLSDGSKDLISKLLRYHPPQRLPLKGVME 339

Qy        380 HPWITANSSK--PSNCQNKES 398
              |||: ||| :  |   |: :|
Db        340 HPWVKANSRRVLPPVYQSTQS 360


RESULT 15
Q6DE08
ID   Q6DE08      PRELIMINARY;      PRT;   361 AA.
AC   Q6DE08;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   LOC398457 protein.
GN   Name=LOC398457;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Klein S., Strausberg R.;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC077339; AAH77339.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   361 AA;  41735 MW;  384803072040015B CRC64;

  Query Match           50.6%;  Score 1067.5;  DB 2;  Length 361;
  Best Local Similarity   55.6%;  Pred. No. 4e-60;
  Matches  212;  Conservative   61;  Mismatches   63;  Indels   45;  Gaps    8;

Qy         35 NPLPVNSGQAQRVLCPSN--SSQRVPLQAQKLVSSH-KPVQNQKQKQLQ--------ATS 83
              ||    |    :   ||:  ::||| |: :  ||:   |  |   :: |        ::|
Db          8 NP----SSYTSKFTTPSSATAAQRV-LRKEPYVSTFTTPSDNLLAQRTQLSRITPSASSS 62

Qy         84 VPHPVS----RPLNNTQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPL 139
              ||  |:     |  ||  :: |                        ||:: ::||:|||||
Db         63 VPGRVAVSTEMPSQNTALAEMP----------------------KRKFTIDDFDIGRPL 99

Qy        140 GKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGY 199
              |||||||||||||||:|||:|||||||:|||| |||||||||:|||||||||||||:| |
Db        100 GKGKFGNVYLAREKQNKFIMALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNY 159

Qy        200 FHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKP 259
              |||  |:||:||:|| | :|:|||| :|||||:||:: |||:|| ||| ::||||||||
Db        160 FHDRKRIYLMLEFAPRGELYKELQKHGRFDEQRSATFMEELADALHYCHERKVIHRDIKP 219

Qy        260 ENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLC 319
              ||||:|  |||||||||||||||| || |:|||||||||||||: ||||||||  ||||
Db        220 ENLLMGYKGELKIADFGWSVHAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLC 279

Qy        320 YEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLE 379
              |||||| |||:: :: ||::||  |:  || |:::|::||||:||:::| ||  |: |:|
Db        280 YEFLVGMPPFDSPSHTETHRRIVNVDLKFPPFLSDGSKDLISKLLRYHPPQRLPLKGVME 339

Qy        380 HPWITANSSK--PSNCQNKES 398
              |||: ||| :  |   |: :|
Db        340 HPWVKANSRRVLPPVYQSTQS 360


RESULT 16
```

```
Q7ZYT9
ID   Q7ZYT9      PRELIMINARY;      PRT;   371 AA.
AC   Q7ZYT9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LOC398457 protein (Fragment).
GN   Name=LOC398457;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC041288; AAH41288.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT   NON_TER       1       1
SQ   SEQUENCE   371 AA;  42716 MW;  19F0C009D0E38CFC CRC64;

  Query Match             50.6%;  Score 1067.5;  DB 2;  Length 371;
  Best Local Similarity   55.6%;  Pred. No. 4.1e-60;
  Matches  212;  Conservative   61;  Mismatches   63;  Indels   45;  Gaps    8;

Qy         35 NPLPVNSGQAQRVLCPSN--SSQRVPLQAQKLVSSH-KPVQNQKQKQLQ--------ATS 83
              ||    |    :   ||:  ::||| |: :  ||:   |  |   :: |        ::|
Db         18 NP----SSYTSKFTTPSSATAAQRV-LRKEPYVSTFTTPSDNLLAQRTQLSRITPSASSS 72

Qy         84 VPHPVS----RPLNNTQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPL 139
              ||  |:    |  ||  :: |                        ||:: ::||:|||||
Db         73 VPGRVAVSTEMPSQNTALAEMP----------------------KRKFTIDDFDIGRPL 109

Qy        140 GKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGY 199
              ||||||||||||||:|||:|||||||:|||| |||||||||:|||||||||||||:| |
Db        110 GKGKFGNVYLAREKQNKFIMALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNY 169

Qy        200 FHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKP 259
              |||  |:||:||:|| | :|:|||| :|||||:||:: |||:|| ||| ::||||||||
Db        170 FHDRKRIYLMLEFAPRGELYKELQKHGRFDEQRSATFMEELADALHYCHERKVIHRDIKP 229

Qy        260 ENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLC 319
              ||||:|  |||||||||||||||| || |:|||||||||||||: ||||||||  ||||
Db        230 ENLLMGYKGELKIADFGWSVHAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLC 289

Qy        320 YEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLE 379
              ||||| |||:: :: ||::||  |:  || |:::|::||||:||:::| ||  |: |:|
Db        290 YEFLVGMPPFDSPSHTETHRRIVNVDLKFPPFLSDGSKDLISKLLRYHPPQRLPLKGVME 349

Qy        380 HPWITANSSK--PSNCQNKES 398
              |||: ||| :  |   |: :|
Db        350 HPWVKANSRRVLPPVYQSTQS 370


RESULT 17
AURB_MOUSE
ID   AURB_MOUSE      STANDARD;      PRT;   345 AA.
AC   O70126; Q61882;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase 12 (EC 2.7.1.37) (Aurora-related
DE   kinase 2) (Serine/threonine-protein kinase 5) (STK-1) (Aurora-B).
GN   Name=Aurkb; Synonyms=Ark2, Stk1, Stk12, Stk5;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Testis;
RX   MEDLINE=96194801; PubMed=8647446; DOI=10.1016/0378-1119(95)00809-8;
RA   Niwa H., Abe K., Kunisada T., Yamamura K.;
RT   "Cell-cycle-dependent expression of the STK-1 gene encoding a novel
RT   murine putative protein kinase.";
RL   Gene 169:197-201(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=98183439; PubMed=9514916; DOI=10.1006/bbrc.1998.8250;
RA   Shindo M., Nakano H., Kuroyanagi H., Shirasawa T., Mihara M.,
RA   Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;
RT   "cDNA cloning, expression, subcellular localization, and chromosomal
RT   assignment of mammalian aurora homologues, aurora-related kinase (ARK)
RT   1 and 2.";
RL   Biochem. Biophys. Res. Commun. 244:285-292(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
```

```
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May be directly involved in regulating the cleavage of
CC       polar spindle microtubules and is a key regulator for the onset of
CC       cytokinesis during mitosis (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localized to the midzone of central spindle
CC       in late anaphase and concentrated into the midbody in telophase
CC       and cytokinesis. Colocalized with gamma tubulin in the mid-body
CC       (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in testis, intestine and spleen. All
CC       of them are tissues that contain a large number of proliferating
CC       cells. Expressed during S phase, in a cell-cycle-dependent
CC       fashion.
CC   -!- DEVELOPMENTAL STAGE: Strongly expressed in 8.5 and 12.5 dpc.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D21099; BAA04658.1; -.
DR   EMBL; U69107; AAC12683.1; -.
DR   EMBL; BC003261; AAH03261.1; -.
DR   HSSP; P31751; 1GZK.
DR   MGD; MGI:107168; Aurkb.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       82    332       Protein kinase.
FT   NP_BIND      88     96       ATP (By similarity).
FT   BINDING     111    111       ATP (By similarity).
FT   ACT_SITE    205    205       Proton acceptor (By similarity).
FT   CONFLICT     44     44       R -> W (in Ref. 1).
SQ   SEQUENCE   345 AA;  39324 MW;  C1D1D8B6BDEA00A0 CRC64;

  Query Match             50.3%;  Score 1061;  DB 1;  Length 345;
  Best Local Similarity   59.0%;  Pred. No. 9.7e-60;
  Matches  209;  Conservative   48;  Mismatches   73;  Indels   24;  Gaps    4;

Qy         36 PLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNT 95
              | |  |  :|  |  :  ||||                :|:  ||     |:|  :|:
Db          9 PWPYGSKTSQSGL--NTLSQRV----------------LRKEPATTSALALVNR--SNS 47

Qy         96 QKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQS 155
              | :  |     ||  :   ||:       |: : :::||||||||||||||||||||:|
Db         48 QSTAAPGQKLAENKSQGSTASQ---GSQNKQPFTIDNFEIGRPLGKGKFGNVYLAREKKS 104

Qy        156 KFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPL 215
```

```
             :||:|||:|||:|:|| |||||||||:|||:||:|||||:|| ||:|  |:||||||||
Db        105 RFIVALKILFKSQIEKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPR 164

Qy        216 GTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADF 275
              | :|:||||   |||||||| : ||::||:||| |:|||||||||||||  ||||||||
Db        165 GELYKELQKSRTFDEQRTATIMEELSDALTYCHKKKVIHRDIKPENLLLGLQGELKIADF 224

Qy        276 GWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQ 335
              |||||||| || |:|||||||||||||||:| |||| :|||||  :|| |||: ::
Db        225 GWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHS 284

Qy        336 ETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
              |||:|| :|:  ||  |  ||:||||:|||||| ||  | ||  |||: ||| :
Db        285 ETYRRIVKVDLKFPSSVPSGAQDLISKLLKHNPWQRLPLAEVAAHPWVRANSRR 338


RESULT 18
AURB_HUMAN
ID   AURB_HUMAN     STANDARD;      PRT;   344 AA.
AC   Q96GD4; O14630; O60446; O95083; Q96DV5; Q9UQ46;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase 12 (EC 2.7.1.37) (Aurora- and Ipl1-
DE   like midbody-associated protein 1) (AIM-1) (Aurora/IPL1-related kinase
DE   2) (Aurora-related kinase 2) (STK-1) (Aurora-B).
GN   Name=AURKB; Synonyms=AIK2, AIM1, ARK2, STK12;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98183439; PubMed=9514916; DOI=10.1006/bbrc.1998.8250;
RA   Shindo M., Nakano H., Kuroyanagi H., Shirasawa T., Mihara M.,
RA   Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;
RT   "cDNA cloning, expression, subcellular localization, and chromosomal
RT   assignment of mammalian aurora homologues, aurora-related kinase (ARK)
RT   1 and 2.";
RL   Biochem. Biophys. Res. Commun. 244:285-292(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99025855; PubMed=9809983;
RA   Tatsuka M., Katayama H., Ota T., Tanaka T., Odashima S., Suzuki F.,
RA   Terada Y.;
RT   "Multinuclearity and increased ploidy caused by overexpression of the
RT   aurora- and Ipl1-like midbody-associated protein mitotic kinase in
RT   human cancer cells.";
RL   Cancer Res. 58:4811-4816(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver, and Spleen;
RX   MEDLINE=99077743; PubMed=9858806;
RA   Kimura M., Matsuda Y., Yoshioka T., Sumi N., Okano Y.;
RT   "Identification and characterization of STK12/Aik2: a human gene
RT   related to aurora of Drosophila and yeast IPL1.";
RL   Cytogenet. Cell Genet. 82:147-152(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21364697; PubMed=11471245;
RA   Prigent C., Gill R., Trower M., Sanseau P.;
RT   "In silico cloning of a new protein kinase, Aik2, related to
RT   Drosophila aurora using the new tool: EST Blast.";
RL   In Silico Biol. 1:123-128(1999).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Zhang Q., Yu L., Bi A.;
RT   "Cloning of a novel human gene homologous to mouse STK-1.";
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung, Lymph, and Muscle;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
```

```
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   REVIEW.
RX   MEDLINE=21306577; PubMed=11413462; DOI=10.1038/35048096;
RA   Nigg E.A.;
RT   "Mitotic kinases as regulators of cell division and its checkpoints.";
RL   Nat. Rev. Mol. Cell Biol. 2:21-32(2001).
CC   -!- FUNCTION: May be directly involved in regulating the cleavage of
CC       polar spindle microtubules and is a key regulator for the onset of
CC       cytokinesis during mitosis.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localized to the midzone of central spindle
CC       in late anaphase and concentrated into the midbody in telophase
CC       and cytokinesis. Colocalized with gamma tubulin in the mid-body.
CC   -!- TISSUE SPECIFICITY: High level expression seen in the thymus. It
CC       is also expressed in the spleen, lung, testis, colon, placenta and
CC       fetal liver. Expressed during S and G2/M phase and expression is
CC       up-regulated in cancer cells during M phase.
CC   -!- DISEASE: Disruptive regulation of expression is a possibile
CC       mechanism of the perturbation of chromosomal integrity in cancer
CC       cells through its dominant-negative effect on cytokinesis.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF008552; AAC12709.1; -.
DR   EMBL; AB011450; BAA32136.1; -.
DR   EMBL; AB011446; BAA82709.1; -.
DR   EMBL; AF004022; AAB65786.1; -.
DR   EMBL; AF015254; AAC98891.1; -.
DR   EMBL; BC000442; AAH00442.2; ALT_INIT.
DR   EMBL; BC009751; AAH09751.1; -.
DR   EMBL; BC013300; AAH13300.2; ALT_INIT.
DR   HSSP; P31751; 1GZK.
DR   Genew; HGNC:11390; AURKB.
DR   H-InvDB; HIX0019005; -.
DR   MIM; 604970; -.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       77    327       Protein kinase.
FT   NP_BIND      83     91       ATP (By similarity).
FT   BINDING     106    106       ATP (By similarity).
FT   ACT_SITE    200    200       Proton acceptor (By similarity).
FT   CONFLICT     14     15       RQ -> DK (in Ref. 5).
FT   CONFLICT     70     70       R -> RR (in Ref. 6; AAH13300).
FT   CONFLICT    161    161       E -> M (in Ref. 4 and 5).
FT   CONFLICT    167    169       QKS -> HKT (in Ref. 4).
FT   CONFLICT    179    179       T -> TVRR (in Ref. 4).
FT   CONFLICT    180    180       I -> VRAV (in Ref. 5).
FT   CONFLICT    226    226       P -> T (in Ref. 3).
FT   CONFLICT    249    250       MH -> ID (in Ref. 3).
FT   CONFLICT    271    271       Missing (in Ref. 3).
FT   CONFLICT    298    298       T -> M (in Ref. 6; AAH09751/AAH13300).
SQ   SEQUENCE   344 AA;  39280 MW;  8325E3EF5A1FB170 CRC64;

  Query Match           50.1%;  Score 1057.5;  DB 1;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 1.6e-59;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |          |:|: |||      :| :   |  ||  |:        | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              ||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 19
Q8C6C1
ID   Q8C6C1      PRELIMINARY;      PRT;   345 AA.
AC   Q8C6C1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus ES cells cDNA, RIKEN full-length enriched library,
DE   clone:2410042H09 product:serine/threonine kinase 5, full insert
DE   sequence.
GN   Name=Aurkb;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
```

```
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
RA   Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AK075951; BAC36078.1; -.
DR   HSSP; O14965; 1OL6.
DR   MGD; MGI:107168; Aurkb.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   345 AA;  39384 MW;  D204D8B91CFF00A4 CRC64;

  Query Match           50.1%;  Score 1057;  DB 2;  Length 345;
  Best Local Similarity  59.0%;  Pred. No. 1.8e-59;

  Matches  209;  Conservative   47;  Mismatches   74;  Indels   24;  Gaps    4;

Qy         36 PLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNT 95
              | | |  :|  |  :  ||||                :|:   ||    |:|   |:
Db          9 PWPYGSKTSQSGL--NTLSQRV----------------LRKEPATTSALALVNR--FNS 47

Qy         96 QKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQS 155
              | :  |     ||  :   ||:       |: : :::||||||||||||||||||||:|
Db         48 QSTAAPGQKLAENKSQGSTASQ---GSQNKQPFTIDNFEIGRPLGKGKFGNVYLAREKKS 104

Qy        156 KFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPL 215
             :||:|||:|||:|:|| |||||||||:|||:||:|||||:|| ||:|  |:||||||||
Db        105 RFIVALKILFKSQIEKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPR 164

Qy        216 GTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADF 275
              | :|:||||   |||||||| : ||::||:||| |:|||||||||||||  ||||||||
Db        165 GELYKELQKSRTFDEQRTATIMEELSDALTYCHKKKVIHRDIKPENLLLGLQGELKIADF 224

Qy        276 GWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQ 335
              |||||||| || |:|||||||||||||||:| |||| :|||||| :|| ||||: ::
Db        225 GWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHS 284

Qy        336 ETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
             |||:|| :|:  ||  |  ||:||||:|||||| ||  | ||  |||: ||| :
Db        285 ETYRRIVKVDLKFPSSVPSGAQDLISKLLKHNPWQRLPLAEVAAHPWVRANSRR 338


RESULT 20
AURB_RAT
ID   AURB_RAT       STANDARD;      PRT;   343 AA.
AC   O55099;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase 12 (EC 2.7.1.37) (Aurora- and Ipl1-
DE   like midbody-associated protein 1) (AIM-1) (Aurora-B).
GN   Name=Aurkb; Synonyms=Aim1, Stk12;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=98119810; PubMed=9450992; DOI=10.1093/emboj/17.3.667;
RA   Terada Y., Tatsuka M., Suzuki F., Yasuda Y., Fujita S., Otsu M.;
RT   "AIM-1: a mammalian midbody-associated protein required for
RT   cytokinesis.";
RL   EMBO J. 17:667-676(1998).
CC   -!- FUNCTION: May be directly involved in regulating the cleavage of
CC       polar spindle microtubules and is a key regulator for the onset of
CC       cytokinesis during mitosis (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localized to the midzone of central spindle
CC       in late anaphase and concentrated into the midbody in telophase
CC       and cytokinesis. Colocalized with gamma tubulin in the mid-body
CC       (By similarity).
CC   -!- TISSUE SPECIFICITY: High level expression seen in the testis. It
CC       is also expressed in the spleen, lung and heart. Expressed in the
CC       G2/M phase of the cell cycle.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D89731; BAA23794.1; -.
DR   HSSP; P31751; 1GZK.
```

```
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       80    330       Protein kinase.
FT   NP_BIND       86     94       ATP (By similarity).
FT   BINDING      109    109       ATP (By similarity).
FT   ACT_SITE     203    203       Proton acceptor (By similarity).
SQ   SEQUENCE   343 AA;  39234 MW;  27B740D20E287598 CRC64;

  Query Match             49.8%;  Score 1050;  DB 1;  Length 343;
  Best Local Similarity   57.6%;  Pred. No. 4.9e-59;
  Matches  204;  Conservative   51;  Mismatches   73;  Indels   26;  Gaps    4;

Qy         36 PLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNT 95
              | |  |  :|  |  :   ||| |: :  |:  : : |:                :|:
Db          9 PWPYGSKTSQSGL--NTLPQRV-LRKEPAVTPAQALMNR-----------------SNS 47

Qy         96 QKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQS 155
              | :  |     ||     |  |      :: : :::|||||||||||||||||||||:|
Db         48 QSTAVPGQKLTENKGATALQGSQ-----SRQPFTIDNFEIGRPLGKGKFGNVYLAREKKS 102

Qy        156 KFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPL 215
             :||:|||:|||:|:|| |||||||||:|||:||:|||||:|| ||:|  |:|||||||
Db        103 RFIVALKILFKSQIEKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPR 162

Qy        216 GTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADF 275
              | :|:|||||   |||||||| : ||::|| ||| |:|||||||||||||  ||||||||
Db        163 GELYKELQKSGTFDEQRTATIMEELSDALMYCHKKKVIHRDIKPENLLLGLQGELKIADF 222

Qy        276 GWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQ 335
             |||||||| || |:|||||||||||||||||:| |||| :|||||| :|| ||||: ::
Db        223 GWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHS 282

Qy        336 ETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
             |||:|| :|:  ||  :  ||:||||:||||||||  | :|  |||: ||| :
Db        283 ETYRRIVKVDLKFPSSMPLGAKDLISKLLKHNPSQRLPLEQVSAHPWVRANSRR 336


RESULT 21
Q7YRC6
ID   Q7YRC6      PRELIMINARY;      PRT;   344 AA.
AC   Q7YRC6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Serine/threonine kinase 12.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G., Li W., Yu L.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY336975; AAQ16151.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   344 AA;  39441 MW;  4DD7158CF2F5D047 CRC64;

  Query Match             49.5%;  Score 1045;  DB 2;  Length 344;
  Best Local Similarity   56.9%;  Pred. No. 1e-58;
  Matches  209;  Conservative   44;  Mismatches   68;  Indels   46;  Gaps    5;

Qy         32 PCQNPLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRP 91
              |  | ||      ||||      :: |:    ||             | : |   | : |
Db         19 PGLNTLP------QRVL------RKEPVTPSALV------------LMSRSNAQPTAAP 53

Qy         92 LNNTQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAR 151
                 ::    |: |                    || : ::||||||||||||||||||||
Db         54 GQKVVENSSGTPNIP-----------------KRSFTIDDFEIGRPLGKGKFGNVYLAR 95

Qy        152 EKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILE 211
             ||:| ||:|||||||:|:|| |||||||||:|||:||:|||||||| ||:|  |:|||||
Db         96 EKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLQHPNILRLYNYFYDRRRIYLILE 155

Qy        212 YAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELK 271
             ||| | :|:|||||   |||||||| : |||:||:|||:|:|||||||||||||  ||||
Db        156 YAPRGELYKELQKSRTFDEQRTATIMEELADALTYCHAKKVIHRDIKPENLLLGLRGELK 215

Qy        272 IADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEA 331
             |||||||||| || |:|||||||||||||||| |:|||||| :|||||| ||| ||||:
Db        216 IADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRTHNEKVDLWCIGVLCYELLVGNPPFES 275

Qy        332 NTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK-- 389
              :: |||:|| :|:  ||  |  ||:|  | :||||||:|  | :|  |||:  :| :
Db        276 ASHNETYRRIVKVDLKFPPSVPLGAQDFIYKLLKHNPSERLPLAQVSAHPWVRTHSRRVL 335

Qy        390 -PSNCQN 395
              ||  |:
Db        336 PPSAPQS 342


Search completed: June 20, 2005, 20:12:24
Job time : 141.649 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:59:50 ; Search time 37.7617 Seconds
(without alignments)
680.035 Million cell updates/sec

Title: US-10-734-126-3
Perfect score: 1809
Sequence: 1 MAQKENSYPWPYGRQTAPSG.....PWVRANSRRVLPPSALQSVA 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 13

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1809	100.0	344	2	US-08-755-728-3	Sequence 3, Appli
2	1809	100.0	344	2	US-08-974-655-3	Sequence 3, Appli
3	1809	100.0	344	3	US-09-283-011-3	Sequence 3, Appli
4	1809	100.0	344	4	US-09-012-135A-3	Sequence 3, Appli
5	1809	100.0	363	4	US-09-949-016-9826	Sequence 9826, Ap
6	1797.5	99.4	347	2	US-09-016-000-1	Sequence 1, Appli
7	1520.5	84.1	343	4	US-09-485-534-4	Sequence 4, Appli
8	1057.5	58.5	403	2	US-08-755-728-4	Sequence 4, Appli
9	1057.5	58.5	403	2	US-08-974-655-4	Sequence 4, Appli
10	1057.5	58.5	403	3	US-09-283-011-4	Sequence 4, Appli
11	1057.5	58.5	403	4	US-09-012-135A-4	Sequence 4, Appli
12	1057.5	58.5	488	4	US-09-949-016-10461	Sequence 10461, A
13	1057.5	58.5	488	4	US-09-949-016-10462	Sequence 10462, A

ALIGNMENTS

RESULT 1
US-08-755-728-3
; Sequence 3, Application US/08755728
; Patent No. 5962312
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Mossie, Kevin
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
; TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,728
; FILING DATE: No. 5962312ember 25, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/008,809
; FILING DATE: December 18, 1995
; APPLICATION NUMBER: 60/023,943
; FILING DATE: August 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 223/113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-755-728-3

Query Match 100.0%; Score 1809; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 4.1e-176;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 2

```
US-08-974-655-3
; Sequence 3, Application US/08974655
; Patent No. 5972676
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,655
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 5972676ember 25, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  223/113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  344 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-974-655-3

  Query Match           100.0%;  Score 1809;  DB 2;  Length 344;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-176;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 3
US-09-283-011-3
; Sequence 3, Application US/09283011
; Patent No. 6207401
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/283,011
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/012,135
;      FILING DATE:  January 22, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 6207401ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  344 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-283-011-3

  Query Match           100.0%;  Score 1809;  DB 3;  Length 344;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-176;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
```

```
Qy          61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy         121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy         181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy         241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy         301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
               |||||||||||||||||||||||||||||||||||||||||||
Db         301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 4
US-09-012-135A-3
; Sequence 3, Application US/09012135A
; Patent No. 6716575
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,135A
;      FILING DATE:  January 22, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/005,268
;      FILING DATE:  January 9, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 6716575ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  344 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-012-135A-3

  Query Match           100.0%;  Score 1809;  DB 4;  Length 344;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-176;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy          61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy         121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy         181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy         241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy         301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
               |||||||||||||||||||||||||||||||||||||||||||
Db         301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 5
US-09-949-016-9826
; Sequence 9826, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9826
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-9826

  Query Match           100.0%;  Score 1809;  DB 4;  Length 363;
  Best Local Similarity  100.0%;  Pred. No. 4.4e-176;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          20 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 79

Qy          61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          80 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 139

Qy         121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         140 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 199
```

```
Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 259

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 319

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        320 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 363


RESULT 6
US-09-016-000-1
; Sequence 1, Application US/09016000
; Patent No. 5962232
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Akerblom, Ingrid E.
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    TITLE OF INVENTION:  PROTEIN KINASE MOLECULES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,000
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0465 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  347 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  HMC1NOT01
;      CLONE:  2940
US-09-016-000-1

  Query Match             99.4%;  Score 1797.5;  DB 2;  Length 347;
  Best Local Similarity   99.1%;  Pred. No. 6.2e-175;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    1;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTAT- 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATV 180

Qy        180 --IMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RAIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 240

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 300

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 347


RESULT 7
US-09-485-534-4
; Sequence 4, Application US/09485534
; Patent No. 6759212
; GENERAL INFORMATION:
;  APPLICANT:  TATSUKA, MASAAKI
;  APPLICANT:  TERADA, YASUHIKO
;  TITLE OF INVENTION:  CELL CYCLE-REGULATING PROTEINS
;  FILE REFERENCE:  050499/0102
;  CURRENT APPLICATION NUMBER: US/09/485,534
;  CURRENT FILING DATE:  2000-02-14
;  PRIOR APPLICATION NUMBER:  PCT/JP98/03641
;  PRIOR FILING DATE:  1998-08-17
;  PRIOR APPLICATION NUMBER:  235371/1997
;  PRIOR FILING DATE:  1997-08-15
;  NUMBER OF SEQ ID NOS:  19
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Rattus sp.
US-09-485-534-4

  Query Match             84.1%;  Score 1520.5;  DB 4;  Length 343;
  Best Local Similarity   84.5%;  Pred. No. 1.1e-146;
  Matches  290;  Conservative   20;  Mismatches   30;  Indels    3;  Gaps    1;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              |||||| |||||| :|: |||:||||||||||||     |  ||:||| | || ||||: ||
Db          1 MAQKENVYPWPYGSKTSQSGLNTLPQRVLRKEPAVTPAQALMNRSNSQSTAVPGQKLTEN 60

Qy         61 SSGTP---DILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEG 117
                 |         : ||||:||||||||||||||||||||| ||||||:||||||||||
Db         61 KGATALQGSQSRQPFTIDNFEIGRPLGKGKFGNVYLAREKKSRFIVALKILFKSQIEKEG 120

Qy        118 VEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRT 177
              |||||||||||||||| ||||||:|||||||::|||||||||||||||||||||| |||||||
Db        121 VEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPRGELYKELQKSGTFDEQRT 180

Qy        178 ATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
              |||||||:||||||| ||||||||||||||||:|||||||||||||||||||||||||||||
Db        181 ATIMEELSDALMYCHKKKVIHRDIKPENLLLGLQGELKIADFGWSVHAPSLRRKTMCGTL 240

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              |||||||||||||| ||||||||||||:|||||||| ||:|||||||||||||:|:|
Db        241 DYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHSETYRRIVKVDLKFPSSMP 300

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              ||:|||||||:||||:|||| ||||||||||||||||||||||
Db        301 LGAKDLISKLLKHNPSQRLPLEQVSAHPWVRANSRRVLPPSAL 343
```

```
RESULT 8
US-08-755-728-4
; Sequence 4, Application US/08755728
; Patent No. 5962312
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/755,728
;      FILING DATE:  No. 5962312ember 25, 1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  223/113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-755-728-4

  Query Match             58.5%;  Score 1057.5;  DB 2;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 2.5e-99;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
                ::|||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              ||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 9
US-08-974-655-4
; Sequence 4, Application US/08974655
; Patent No. 5972676
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,655
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 5972676ember 25, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  223/113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-974-655-4

  Query Match             58.5%;  Score 1057.5;  DB 2;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 2.5e-99;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
                ::|||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:|||:||
```

```
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||| || |:||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 10
US-09-283-011-4
; Sequence 4, Application US/09283011
; Patent No. 6207401
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/283,011
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/012,135
;      FILING DATE:  January 22, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 6207401ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO

US-09-283-011-4

  Query Match           58.5%;  Score 1057.5;  DB 3;  Length 403;
  Best Local Similarity  63.2%;  Pred. No. 2.5e-99;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||| || |:||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 11
US-09-012-135A-4
; Sequence 4, Application US/09012135A
; Patent No. 6716575
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,135A
;      FILING DATE:  January 22, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/005,268
;      FILING DATE:  January 9, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. 6716575ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-012-135A-4

  Query Match           58.5%;  Score 1057.5;  DB 4;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 2.5e-99;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db        82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              ||||||||  ||:|  |:||||||||  | :|:||||   ||||||||  : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  ||||||||||||||| || |:|||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              ||||  :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :| | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 12
US-09-949-016-10461
; Sequence 10461, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10461
;   LENGTH: 488
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10461

  Query Match           58.5%;  Score 1057.5;  DB 4;  Length 488;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-99;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db        82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              ||||||||  ||:|  |:||||||||  | :|:||||   ||||||||  : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  ||||||||||||||| || |:|||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              ||||  :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :| | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 13
US-09-949-016-10462
; Sequence 10462, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10462
;   LENGTH: 488
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10462

  Query Match           58.5%;  Score 1057.5;  DB 4;  Length 488;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-99;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db        167 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 214

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        215 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 274

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              ||||||||  ||:|  |:||||||||  | :|:||||   ||||||||  : |||:|| |||
Db        275 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 334

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  ||||||||||||||| || |:|||||||||||||||||:||
Db        335 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 394

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              ||||  :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        395 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 454

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :| | :|  |||: ||| :   ||  |:
Db        455 QRPMLREVLEHPWITANSSK---PSNCQN 480
```

```
Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        455 QRPMLREVLEHPWITANSSK---PSNCQN 480
```

Search completed: June 20, 2005, 20:13:50
Job time : 41.7617 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:47:43 ; Search time 66.3132 Seconds
(without alignments)
1991.894 Million cell updates/sec

Title: US-10-734-126-3
Perfect score: 1809
Sequence: 1 MAQKENSYPWPYGRQTAPSG.....PWVRANSRRVLPPSALQSVA 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 49

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1809	100.0	344	9	US-09-012-135A-3	Sequence 3, Appli
2	1809	100.0	344	11	US-09-784-332-3	Sequence 3, Appli
3	1809	100.0	344	14	US-10-060-065-13	Sequence 13, Appl
4	1809	100.0	344	14	US-10-059-585-34	Sequence 34, Appl
5	1809	100.0	344	14	US-10-171-311-214	Sequence 214, App
6	1809	100.0	344	15	US-10-295-027-203	Sequence 203, App
7	1809	100.0	344	15	US-10-173-999-101	Sequence 101, App
8	1809	100.0	344	15	US-10-188-832-93	Sequence 93, Appl
9	1809	100.0	344	16	US-10-734-126-3	Sequence 3, Appli
10	1797.5	99.4	347	9	US-09-974-298-136	Sequence 136, App
11	1520.5	84.1	343	15	US-10-429-849-4	Sequence 4, Appli
12	1181	65.3	355	15	US-10-092-900A-14	Sequence 14, Appl
13	1124	62.1	320	15	US-10-403-571-98	Sequence 98, Appl
14	1057.5	58.5	322	16	US-10-664-421-125	Sequence 125, App
15	1057.5	58.5	403	9	US-09-012-135A-4	Sequence 4, Appli
16	1057.5	58.5	403	11	US-09-784-332-4	Sequence 4, Appli
17	1057.5	58.5	403	14	US-10-060-065-12	Sequence 12, Appl
18	1057.5	58.5	403	14	US-10-059-585-33	Sequence 33, Appl
19	1057.5	58.5	403	14	US-10-209-324-2	Sequence 2, Appli
20	1057.5	58.5	403	15	US-10-295-027-1206	Sequence 1206, Ap
21	1057.5	58.5	403	15	US-10-087-684-47	Sequence 47, Appl
22	1057.5	58.5	403	15	US-10-087-684-48	Sequence 48, Appl
23	1057.5	58.5	403	15	US-10-087-684-49	Sequence 49, Appl
24	1057.5	58.5	403	15	US-10-218-779-47	Sequence 47, Appl
25	1057.5	58.5	403	15	US-10-218-779-48	Sequence 48, Appl
26	1057.5	58.5	403	15	US-10-218-779-49	Sequence 49, Appl
27	1057.5	58.5	403	15	US-10-188-832-35	Sequence 35, Appl
28	1057.5	58.5	403	16	US-10-751-736-65	Sequence 65, Appl
29	1057.5	58.5	403	16	US-10-734-126-4	Sequence 4, Appli
30	1057.5	58.5	403	17	US-10-781-581-187	Sequence 187, App
31	1057.5	58.5	420	15	US-10-264-049-3090	Sequence 3090, Ap
32	1050.5	58.1	403	14	US-10-026-021-7	Sequence 7, Appli
33	1050.5	58.1	403	15	US-10-087-684-50	Sequence 50, Appl
34	1050.5	58.1	403	15	US-10-218-779-50	Sequence 50, Appl
35	1045.5	57.8	402	15	US-10-087-684-51	Sequence 51, Appl
36	1045.5	57.8	402	15	US-10-218-779-51	Sequence 51, Appl
37	966	53.4	403	15	US-10-087-684-14	Sequence 14, Appl
38	966	53.4	403	15	US-10-218-779-14	Sequence 14, Appl
39	940	52.0	294	15	US-10-389-566-1246	Sequence 1246, Ap
40	937	51.8	294	15	US-10-389-566-1606	Sequence 1606, Ap
41	934.5	51.7	295	15	US-10-389-566-645	Sequence 645, App
42	933.5	51.6	298	15	US-10-389-566-644	Sequence 644, App
43	921.5	50.9	282	15	US-10-389-566-1240	Sequence 1240, Ap
44	918	50.7	368	15	US-10-389-566-382	Sequence 382, App
45	917.5	50.7	418	15	US-10-389-566-819	Sequence 819, App
46	916.5	50.7	320	16	US-10-437-963-133574	Sequence 133574,
47	909	50.2	307	15	US-10-424-599-274736	Sequence 274736,
48	908	50.2	298	15	US-10-389-566-550	Sequence 550, App
49	906	50.1	432	15	US-10-389-566-2420	Sequence 2420, Ap

ALIGNMENTS

RESULT 1
US-09-012-135A-3
; Sequence 3, Application US/09012135A
; Patent No. US20020081578A1
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Mossie, Kevin
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
; TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,135A
; FILING DATE: January 22, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/005,268

```
;       FILING DATE:  January 9, 1998
;       APPLICATION NUMBER:  08/755,728
;       FILING DATE:  No. US20020081578A1ember 25, 1996
;       APPLICATION NUMBER:  60/023,943
;       FILING DATE:  August 14, 1996
;       APPLICATION NUMBER:  60/008,809
;       FILING DATE:  December 18, 1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Warburg, Richard J.
;       REGISTRATION NUMBER:  32,327
;       REFERENCE/DOCKET NUMBER:  231/282
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (213) 489-1600
;       TELEFAX:  (213) 955-0440
;       TELEX:  67-3510
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  344 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-09-012-135A-3

  Query Match             100.0%;  Score 1809;  DB 9;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-140;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 2
US-09-784-332-3
; Sequence 3, Application US/09784332
; Publication No. US20050002938A1
;   GENERAL INFORMATION:
;       APPLICANT: Plowman, Gregory
;                  Mossie, Kevin
;       TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
;                           AND/OR AUR-2 RELATED DISORDERS
;       NUMBER OF SEQUENCES: 39
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street
;                    Suite 4700
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: U.S.A.
;            ZIP: 90071-2066
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                         storage
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: IBM P.C. DOS 5.0
;            SOFTWARE: FastSEQ for Windows 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/784,332
;            FILING DATE: 16-Feb-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/283,011
;            FILING DATE: <Unknown>
;            APPLICATION NUMBER: 09/012,135
;            FILING DATE: January 22, 1998
;            APPLICATION NUMBER: 08/755,728
;            FILING DATE: November 25, 1996
;            APPLICATION NUMBER: 60/023,943
;            FILING DATE: August 14, 1996
;            APPLICATION NUMBER: 60/008,809
;            FILING DATE: December 18, 1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Warburg, Richard J.
;            REGISTRATION NUMBER: 32,327
;            REFERENCE/DOCKET NUMBER: 231/282
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (213) 489-1600
;            TELEFAX: (213) 955-0440
;            TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 344 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-784-332-3

  Query Match             100.0%;  Score 1809;  DB 11;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-140;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 3
US-10-060-065-13
; Sequence 13, Application US/10060065
```

; Publication No. US20030017480A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-Ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-099002
; CURRENT APPLICATION NUMBER: US/10/060,065
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-065-13

Query Match 100.0%; Score 1809; DB 14; Length 344;
Best Local Similarity 100.0%; Pred. No. 7.6e-140;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 4
US-10-059-585-34

; Sequence 34, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru
; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Wakamatsu, Ai
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Funahashi, Shin-Ichi
; APPLICANT: Senoo, Chiaki
; APPLICANT: Nezu, Jun-Ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-059-585-34

Query Match 100.0%; Score 1809; DB 14; Length 344;
Best Local Similarity 100.0%; Pred. No. 7.6e-140;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 5
US-10-171-311-214

; Sequence 214, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 214
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-214

Query Match 100.0%; Score 1809; DB 14; Length 344;
Best Local Similarity 100.0%; Pred. No. 7.6e-140;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 6
US-10-295-027-203
; Sequence 203, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 203
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-203

Query Match 100.0%; Score 1809; DB 15; Length 344;
Best Local Similarity 100.0%; Pred. No. 7.6e-140;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 7
US-10-173-999-101
; Sequence 101, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.

```
;  TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
;  TITLE OF INVENTION:  and Methods of Screening for Modulators of Ovarian
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 018501-002420US
;  CURRENT APPLICATION NUMBER: US/10/173,999
;  CURRENT FILING DATE:  2002-06-17
;  PRIOR APPLICATION NUMBER: US 60/299,234
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: US 60/315,287
;  PRIOR FILING DATE: 2001-08-27
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/372,246
;  PRIOR FILING DATE: 2001-04-12
;  NUMBER OF SEQ ID NOS: 163
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-173-999-101

  Query Match             100.0%;  Score 1809;  DB 15;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-140;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 8
US-10-188-832-93
; Sequence 93, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David H.
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
;  TITLE OF INVENTION:  and Methods of Screening for Modulators of Bladder
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 018501-002330US
;  CURRENT APPLICATION NUMBER: US/10/188,832
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: US 60/302,814
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: US 60/310,099
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: US 60/343,705
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/372,246
;  PRIOR FILING DATE: 2002-04-12
;  NUMBER OF SEQ ID NOS: 207
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
;   LENGTH: 344
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-188-832-93

  Query Match             100.0%;  Score 1809;  DB 15;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-140;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 9
US-10-734-126-3
; Sequence 3, Application US/10734126
; Publication No. US20040265852A1
;   GENERAL INFORMATION:
;        APPLICANT: Plowman, Gregory
;                   Mossie, Kevin
;        TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
;                            AND/OR AUR-2 RELATED DISORDERS
;        NUMBER OF SEQUENCES: 39
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/734,126
;             FILING DATE: 15-Dec-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/784,332
;             FILING DATE: 16-Feb-2001
;             APPLICATION NUMBER: 09/283,011
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 09/012,135
```

```
;             FILING DATE: January 22, 1998
;             APPLICATION NUMBER: 08/755,728
;             FILING DATE: November 25, 1996
;             APPLICATION NUMBER: 60/023,943
;             FILING DATE: August 14, 1996
;             APPLICATION NUMBER: 60/008,809
;             FILING DATE: December 18, 1995
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Warburg, Richard J.
;             REGISTRATION NUMBER: 32,327
;             REFERENCE/DOCKET NUMBER: 231/282
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 344 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-734-126-3

  Query Match             100.0%;  Score 1809;  DB 16;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-140;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 10
US-09-974-298-136
; Sequence 136, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, Huei-Mei
;  TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
;  FILE REFERENCE: PA-0037 P
;  CURRENT APPLICATION NUMBER: US/09/974,298
;  CURRENT FILING DATE:  2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,331
;  PRIOR FILING DATE: 2000-05-10
;  NUMBER OF SEQ ID NOS: 194
;  SOFTWARE: PERL Program
; SEQ ID NO 136
;   LENGTH: 347
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020156263A1 002940CD1
US-09-974-298-136

  Query Match             99.4%;  Score 1797.5;  DB 9;  Length 347;
  Best Local Similarity   99.1%;  Pred. No. 6.7e-139;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    1;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTAT- 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATV 180

Qy        180 --IMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RAIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 240

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 300

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 347


RESULT 11
US-10-429-849-4
; Sequence 4, Application US/10429849
; Publication No. US20040029157A1
; GENERAL INFORMATION:
;  APPLICANT: TATSUKA, MASAAKI
;  APPLICANT:  TERADA, YASUHIKO
;  TITLE OF INVENTION: CELL CYCLE-REGULATING PROTEINS
;  FILE REFERENCE: 050499/0102
;  CURRENT APPLICATION NUMBER: US/10/429,849
;  CURRENT FILING DATE:  2003-05-06
;  PRIOR APPLICATION NUMBER: US/09/485,534
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: PCT/JP98/03641
;  PRIOR FILING DATE: 1998-08-17
;  PRIOR APPLICATION NUMBER: 235371/1997
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Rattus sp.
US-10-429-849-4

  Query Match             84.1%;  Score 1520.5;  DB 15;  Length 343;
  Best Local Similarity   84.5%;  Pred. No. 3.2e-116;
  Matches  290;  Conservative   20;  Mismatches   30;  Indels    3;  Gaps    1;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              |||||| |||||| :|: |||:|||||||||||    |  ||:||| | || ||||: ||
Db          1 MAQKENVYPWPYGSKTSQSGLNTLPQRVLRKEPAVTPAQALMNRSNSQSTAVPGQKLTEN 60

Qy         61 SSGTP---DILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEG 117
                 |       : ||||:||||||||||||||||||| ||||||:||||||||||
Db         61 KGATALQGSQSRQPFTIDNFEIGRPLGKGKFGNVYLAREKKSRFIVALKILFKSQIEKEG 120

Qy        118 VEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRT 177
```

```
                ||||||||||||||| |||||:|||||||::|||||||||||||||||||||| |||||||
Db        121 VEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPRGELYKELQKSGTFDEQRT 180

Qy        178 ATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
                ||||||:||||||| |||||||||||||||:|||||||||||||||||||||||||
Db        181 ATIMEELSDALMYCHKKKVIHRDIKPENLLLGLQGELKIADFGWSVHAPSLRRKTMCGTL 240

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
                ||||||||||||||| |||||||||||:|||||||| ||:|||||||||||||:|:|
Db        241 DYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHSETYRRIVKVDLKFPSSMP 300

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
                ||:|||||||:||||:|||| ||||||||||||||||||||
Db        301 LGAKDLISKLLKHNPSQRLPLEQVSAHPWVRANSRRVLPPSAL 343


RESULT 12
US-10-092-900A-14
; Sequence 14, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Taupier Jr., Raymond J.
;  APPLICANT:  Pena, Carol E.A.
;  APPLICANT:  Li, Li
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Gusev, Vladimir Y.
;  APPLICANT:  Ji, Weizhen
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Guo, Xiaojia Sasha
;  APPLICANT:  Tchernev,, Velizar T.
;  APPLICANT:  Fernandes, Elma R.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Liu, Yi
;  APPLICANT:  Anderson, David W.
;  APPLICANT:  Spaderna, Steven K.
;  APPLICANT:  Catterton, Elina
;  APPLICANT:  Leite, Mario W.
;  APPLICANT:  Zhong, Haihong
;  APPLICANT:  Alsobrook, John P.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Catherine E.
;  TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-290C
;  CURRENT APPLICATION NUMBER: US/10/092,900A
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: USSN 60/274,322
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/283,675
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: USSN 60/338,092
;  PRIOR FILING DATE: 2001-12-03
;  PRIOR APPLICATION NUMBER: USSN 60/274,281
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/274,191
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/325,681
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: USSN 60/304,354
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: USSN 60/279,995
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: USSN 60/294,899
;  PRIOR FILING DATE: 2001-05-31
;  PRIOR APPLICATION NUMBER: USSN 60/287,424
;  PRIOR FILING DATE: 2001-04-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 14
;   LENGTH: 355
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-900A-14

  Query Match             65.3%;  Score 1181;  DB 15;  Length 355;
  Best Local Similarity   70.8%;  Pred. No. 2.2e-88;
  Matches  247;  Conservative   27;  Mismatches   61;  Indels   14;  Gaps    6;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              |||||||||||||:||||:|||||  ||| : | | :|  | | :: || |||| :|:||
Db          1 MAQKENSYPWPYGKQTAPAGLSTLLPRVLPRIP-TEAARELPSCADPQPAAAPGHEVVEN 59

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              | |   |||| | :|| | |||||| || :|||||:| ||||||||  |||||| |||||
Db         60 SCGKRSILTRPFLVDDLETGRPLGKDKFVHVYLARKKTSHFIVALKA-FKSQIE-EGVEH 117

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPR----GELYKELQKSCTFDEQR 176
              |:||::||||   ||||| ||||||| |:|| ||||||       |||:||:|| |||::
Db        118 QMRRQMEIQAPFQHPNILSLYNYFYDLRKIYWILEYAPATPTPEELYQELRKSRTFDKKP 177

Qy        177 TATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKT--MC 234
              ||||  |:||||||||||||  ||:||:||| ||:||||:||||  ||||||||||  ||
Db        178 TATITGEVADALMYCHGKKVTPRDMKPDNLLSGLEGELKVADFGCPVHAPSLRRKTRQMC 237

Qy        235 GTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPA 294
              |||||| || |||| | |||||| || | || |||||     :||| | ||||| ||||
Db        238 GTLDYLSPETIEGRAHTEKVDLWYIGALGYEPLVGNP-----THNEAYGRIVKVALKFPL 292

Qy        295 SVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
                |   |||||||||||||||||||||||||| : |:|||||||| |||
Db        293 LCPGEPQDLISKLLRHNPSERLPLAQVSAHPGILAHSRRVLPPSAHQSV 341


RESULT 13
US-10-403-571-98
; Sequence 98, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
;  APPLICANT: Hopkins, Nancy
;  APPLICANT:  Golling, Gregory
;  APPLICANT:  Amsterdam, Adam
;  APPLICANT:  Sun, Zhoaxia
;  TITLE OF INVENTION: Developmental Mutations in Zebrafish
;  FILE REFERENCE: 01997/539002
;  CURRENT APPLICATION NUMBER: US/10/403,571
;  CURRENT FILING DATE:  2003-03-25
;  PRIOR APPLICATION NUMBER: US 60/368,760
;  PRIOR FILING DATE: 2002-03-29
;  NUMBER OF SEQ ID NOS: 159
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 98
;   LENGTH: 320
;   TYPE: PRT
;   ORGANISM: Danio rerio
US-10-403-571-98

  Query Match             62.1%;  Score 1124;  DB 15;  Length 320;
  Best Local Similarity   65.7%;  Pred. No. 9e-84;
  Matches  224;  Conservative   33;  Mismatches   52;  Indels   32;  Gaps    4;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQP----TAAPGQK 56
              |  |||  |                 || :    |||| |    |  : |     : ||:
Db          1 MQNKENREP-----------------RVQQ----TPSAGVGPLRVEMNPDTHAVSGPGRV 39

Qy         57 VMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKE 116
                ::::|       :  :||||:||||||||||||||||:|   ::|||||||||: ||
```

```
Db         40 PVKSNS-------KVLSIDDFDIGRPLGKGKFGNVYLARERKLKVVIALKVLFKSQMVKE 92

Qy        117 GVEHQLRRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQR 176
              |||||||||||||:|| |||||| ||||:|  |::||||||||||:|||||:   ||:||
Db         93 GVEHQLRRREIEIQSHLRHPNILRFYNYFHDDTRVFLILEYAPRGEMYKELQRYGHFDDQR 152

Qy        177 TATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGT 236
              ||| |||::||| ||| |||||||||||||||| :|||||||||||||||||:|||||
Db        153 TATYMEEVSDALQYCHEKKVIHRDIKPENLLLGYRGELKIADFGWSVHAPSLRRRTMCGT 212

Qy        237 LDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASV 296
              ||||||||||  |:|||||| |||||| ||||||||:  | |||:|| ||||:||  |
Db        213 LDYLPPEMIEGHSHDEKVDLWSIGVLCYECLVGNPPFETRQHAETYKRITKVDLQFPKLV 272

Qy        297 PTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
                ||:|||||||||:|| ||||  |  |  |:|||||||||
Db        273 SEGARDLISKLLRHSPSMRLPLRSVMEHRGVKANSRRVLPP 313


RESULT 14
US-10-664-421-125
; Sequence 125, Application US/10664421
; Publication No. US20040142864A1
; GENERAL INFORMATION:
;  APPLICANT: BREMER, RYAN
;  APPLICANT:  IBRAHIM, PRABHA
;  APPLICANT:  KUMAR, ABHINAV
;  APPLICANT:  MANDIYAN, VALSAN
;  APPLICANT:  MILBURN, MICHAEL V.
;  TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE
;  FILE REFERENCE: 039363/0703
;  CURRENT APPLICATION NUMBER: US/10/664,421
;  CURRENT FILING DATE:  2003-09-16
;  PRIOR APPLICATION NUMBER: 60/412,341
;  PRIOR FILING DATE: 2002-09-20
;  PRIOR APPLICATION NUMBER: 60/411,398
;  PRIOR FILING DATE: 2002-09-16
;  NUMBER OF SEQ ID NOS: 169
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 125
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-664-421-125

  Query Match           58.5%;  Score 1057.5;  DB 16;  Length 322;
  Best Local Similarity   63.2%;  Pred. No. 2.5e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||     :| :   |  ||  |:          | :
Db          1 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 48

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         49 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 108

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||||   |||||||| : |||:|| |||
Db        109 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 168

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        169 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 228

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        229 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 288

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        289 QRPMLREVLEHPWITANSSK---PSNCQN 314
```

```
RESULT 15
US-09-012-135A-4
; Sequence 4, Application US/09012135A
; Patent No. US20020081578A1
;  GENERAL INFORMATION:
;    APPLICANT:  Plowman, Gregory
;    APPLICANT:  Mossie, Kevin
;    TITLE OF INVENTION:  DIAGNOSIS AND TREATMENT OF AUR-1
;    TITLE OF INVENTION:  AND/OR AUR-2 RELATED DISORDERS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSEQ for Windows 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,135A
;      FILING DATE:  January 22, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/005,268
;      FILING DATE:  January 9, 1998
;      APPLICATION NUMBER:  08/755,728
;      FILING DATE:  No. US20020081578A1ember 25, 1996
;      APPLICATION NUMBER:  60/023,943
;      FILING DATE:  August 14, 1996
;      APPLICATION NUMBER:  60/008,809
;      FILING DATE:  December 18, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  231/282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  403 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-09-012-135A-4

  Query Match           58.5%;  Score 1057.5;  DB 9;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||     :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
```

```
Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:|||| 
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 16
US-09-784-332-4
; Sequence 4, Application US/09784332
; Publication No. US20050002938A1
;   GENERAL INFORMATION:
;        APPLICANT: Plowman, Gregory
;                   Mossie, Kevin
;        TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
;                            AND/OR AUR-2 RELATED DISORDERS
;        NUMBER OF SEQUENCES: 39
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/784,332
;             FILING DATE: 16-Feb-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/283,011
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 09/012,135
;             FILING DATE: January 22, 1998
;             APPLICATION NUMBER: 08/755,728
;             FILING DATE: November 25, 1996
;             APPLICATION NUMBER: 60/023,943
;             FILING DATE: August 14, 1996
;             APPLICATION NUMBER: 60/008,809
;             FILING DATE: December 18, 1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Warburg, Richard J.
;             REGISTRATION NUMBER: 32,327
;             REFERENCE/DOCKET NUMBER: 231/282
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 403 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-784-332-4

  Query Match           58.5%;  Score 1057.5;  DB 11;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||       :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:|||| 
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 17
US-10-060-065-12
; Sequence 12, Application US/10060065
; Publication No. US20030017480A1
; GENERAL INFORMATION:
;  APPLICANT: Toshio Ota
;  APPLICANT:  Takao Isogai
;  APPLICANT:  Tetsuo Nishikawa
;  APPLICANT:  Koji Hayashi
;  APPLICANT:  Kaoru Otsuka
;  APPLICANT:  Jun-Ichi Yamamoto
;  APPLICANT:  Shizuko Ishii
;  APPLICANT:  Tomoyasu Sugiyama
;  APPLICANT:  Ai Wakamatsu
;  APPLICANT:  Keiichi Nagai
;  APPLICANT:  Tetsuji Otsuki
;  APPLICANT:  Shin-Ichi Funahashi
;  APPLICANT:  Chiaki Senoo
;  APPLICANT:  Jun-Ichi Nezu
;  TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE
;  FILE REFERENCE: 06501-099002
;  CURRENT APPLICATION NUMBER: US/10/060,065
;  CURRENT FILING DATE:  2002-01-29
;  PRIOR APPLICATION NUMBER: PCT/JP00/05061
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: US 60/159,590
;  PRIOR FILING DATE: 1999-10-18
;  PRIOR APPLICATION NUMBER: US 60/183,322
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: JP 11-248036
;  PRIOR FILING DATE: 1999-07-29
;  PRIOR APPLICATION NUMBER: JP 2000-118776
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: JP 2000-183767
;  PRIOR FILING DATE: 2000-05-02
;  PRIOR APPLICATION NUMBER: JP 2000-241899
;  PRIOR FILING DATE: 2000-06-09
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-060-065-12
```

Query Match 58.5%; Score 1057.5; DB 14; Length 403;
Best Local Similarity 63.2%; Pred. No. 3.3e-78;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||    :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
```

RESULT 18
US-10-059-585-33
; Sequence 33, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru
; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Wakamatsu, Ai
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Funahashi, Shin-Ichi
; APPLICANT: Senoo, Chiaki
; APPLICANT: Nezu, Jun-Ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-059-585-33

Query Match 58.5%; Score 1057.5; DB 14; Length 403;
Best Local Similarity 63.2%; Pred. No. 3.3e-78;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||    :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
```

RESULT 19
US-10-209-324-2
; Sequence 2, Application US/10209324
; Publication No. US20030108910A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF CALIFORNIA SAN FRANCISCO
; APPLICANT: TOLAND, Amanda E.
; APPLICANT: BALMAIN, Allan
; TITLE OF INVENTION: STK15 (STK6) GENE POLYMORPHISM AND METHODS OF DETERMINING CANCER P
; FILE REFERENCE: UCSF1120-2
; CURRENT APPLICATION NUMBER: US/10/209,324
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US 60/334,146
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 60/308,911
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (31)..(31)
; OTHER INFORMATION: Xaa is Ile or Phe
US-10-209-324-2

Query Match 58.5%; Score 1057.5; DB 14; Length 403;
Best Local Similarity 63.2%; Pred. No. 3.3e-78;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||    :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
```

```
Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 20
US-10-295-027-1206
; Sequence 1206, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1206
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-295-027-1206

  Query Match            58.5%;  Score 1057.5;  DB 15;  Length 403;
  Best Local Similarity  63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| ||||||||:|||:|||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 21
US-10-087-684-47
; Sequence 47, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 47
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-47

  Query Match            58.5%;  Score 1057.5;  DB 15;  Length 403;
  Best Local Similarity  63.2%;  Pred. No. 3.3e-78;
```

```
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||:| ||:|||||||:|:|| ||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 22
US-10-087-684-48
; Sequence 48, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 48
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-48

  Query Match             58.5%;  Score 1057.5;  DB 15;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||:| ||:|||||||:|:|| ||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 23
US-10-087-684-49
; Sequence 49, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
```

```
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 49
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-49

  Query Match             58.5%;  Score 1057.5;  DB 15;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| ||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||  ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 24
US-10-218-779-47
; Sequence 47, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-47

  Query Match             58.5%;  Score 1057.5;  DB 15;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| ||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||  ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 25
US-10-218-779-48
; Sequence 48, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
```

```
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-48

  Query Match             58.5%;  Score 1057.5;  DB 15;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||   |  ||:||||:||:|||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 26
US-10-218-779-49
; Sequence 49, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-49

  Query Match             58.5%;  Score 1057.5;  DB 15;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||   |  ||:||||:||:|||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
```

```
RESULT 27
US-10-188-832-35
; Sequence 35, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David H.
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
;  TITLE OF INVENTION:  and Methods of Screening for Modulators of Bladder
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 018501-002330US
;  CURRENT APPLICATION NUMBER: US/10/188,832
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: US 60/302,814
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: US 60/310,099
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: US 60/343,705
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/372,246
;  PRIOR FILING DATE: 2002-04-12
;  NUMBER OF SEQ ID NOS: 207
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-188-832-35

  Query Match           58.5%;  Score 1057.5;  DB 15;  Length 403;
  Best Local Similarity  63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||  : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 28
US-10-751-736-65
; Sequence 65, Application US/10751736
; Publication No. US20040265230A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Martinez, Robert
;  APPLICANT:  Brown, Eugene
;  APPLICANT:  Liu, Wei
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
;  TITLE OF INVENTION:  CANCERS
;  FILE REFERENCE: AM100927 (031896-002000)
;  CURRENT APPLICATION NUMBER: US/10/751,736
;  CURRENT FILING DATE:  2003-01-06
;  PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
;  PRIOR FILING DATE: 2003-01-06
;  NUMBER OF SEQ ID NOS: 54873
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 65
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-751-736-65

  Query Match           58.5%;  Score 1057.5;  DB 16;  Length 403;
  Best Local Similarity  63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||  : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 29
US-10-734-126-4
; Sequence 4, Application US/10734126
; Publication No. US20040265852A1
;   GENERAL INFORMATION:
;        APPLICANT: Plowman, Gregory
;                   Mossie, Kevin
;        TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
;                            AND/OR AUR-2 RELATED DISORDERS
;        NUMBER OF SEQUENCES: 39
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/734,126
;             FILING DATE: 15-Dec-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/784,332
```

```
;             FILING DATE: 16-Feb-2001
;             APPLICATION NUMBER: 09/283,011
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 09/012,135
;             FILING DATE: January 22, 1998
;             APPLICATION NUMBER: 08/755,728
;             FILING DATE: November 25, 1996
;             APPLICATION NUMBER: 60/023,943
;             FILING DATE: August 14, 1996
;             APPLICATION NUMBER: 60/008,809
;             FILING DATE: December 18, 1995
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Warburg, Richard J.
;             REGISTRATION NUMBER: 32,327
;             REFERENCE/DOCKET NUMBER: 231/282
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 403 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       ANTI-SENSE: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-734-126-4

  Query Match             58.5%;  Score 1057.5;  DB 16;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||   |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 30
US-10-781-581-187
; Sequence 187, Application US/10781581
; Publication No. US20050019746A1
; GENERAL INFORMATION:
;  APPLICANT: EiRx Therapeutics Ltd.
;  APPLICANT:  Seery, Liam
;  APPLICANT:  Hayes, Ian
;  APPLICANT:  Murphy, Finbarr
;  TITLE OF INVENTION: Apoptosis-Related Kinase/GPCRs
;  FILE REFERENCE: 8912/2015
;  CURRENT APPLICATION NUMBER: US/10/781,581
;  CURRENT FILING DATE:  2004-02-18
;  PRIOR APPLICATION NUMBER: US 10/764,238
;  PRIOR FILING DATE: 2004-01-23
;  PRIOR APPLICATION NUMBER: US 60/457,533
;  PRIOR FILING DATE: 2003-03-25
;  PRIOR APPLICATION NUMBER: UK 0301566.6
;  PRIOR FILING DATE: 2003-01-23
;  NUMBER OF SEQ ID NOS: 226
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 187
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-781-581-187

  Query Match             58.5%;  Score 1057.5;  DB 17;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||   |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 31
US-10-264-049-3090
; Sequence 3090, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3090
;   LENGTH: 420
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-049-3090

  Query Match             58.5%;  Score 1057.5;  DB 15;  Length 420;
  Best Local Similarity   63.2%;  Pred. No. 3.5e-78;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db         99 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 146
```

```
Qy          74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
               ::||||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         147 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 206

Qy         134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
               |||||||| ||:|  |:||||||| | :|:|||   |||||||  : |||:|| |||
Db         207 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 266

Qy         194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
               |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db         267 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 326

Qy         254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
               |||| :|||||  ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:|||
Db         327 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 386

Qy         314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
               :|  | :|  |||: ||| :   ||  |:
Db         387 QRPMLREVLEHPWITANSSK---PSNCQN 412


RESULT 32
US-10-026-021-7
; Sequence 7, Application US/10026021
; Publication No. US20030027756A1
; GENERAL INFORMATION:
;  APPLICANT: Hitoshi, Yasumichi
;  APPLICANT:  Demo, Susan
;  APPLICANT:  Jenkins, Yonchu
;  APPLICANT:  Rigel Pharmaceuticals, Inc.
;  TITLE OF INVENTION: SAK:  Modulation of Cellular Proliferation for
;  TITLE OF INVENTION:  Treatment of Cancer
;  FILE REFERENCE: 021044-001210US
;  CURRENT APPLICATION NUMBER: US/10/026,021
;  CURRENT FILING DATE:  2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/309,632
;  PRIOR FILING DATE: 2001-08-01
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human ARK mitotic kinase
US-10-026-021-7

  Query Match            58.1%;  Score 1050.5;  DB 14;  Length 403;
  Best Local Similarity   61.7%;  Pred. No. 1.2e-77;
  Matches  209;  Conservative   40;  Mismatches   67;  Indels   23;  Gaps    5;

Qy           4 KENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSSG 63
               :: | | |  |          |      |:|: |||      :| :   |  ||  |:
Db          80 QQTSVPHPVSR--------PLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK- 125

Qy          64 TPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
                 | : ::|||||||||||||||||||:|  |:|||||:|:|| |||||||
Db         126 ------RQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLR 179

Qy         124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
               ||:|||:|| |||||||| ||:|  |:||||||| | :|:|||   |||||||  : |
Db         180 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITE 239

Qy         184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
               ||:|| ||| |:|||||||||||||  |||||||||||||||| || |:||||||||||
Db         240 LANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPE 299

Qy         244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
               |||||||:|||||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||
Db         300 MIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDL 359

Qy         304 ISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
               ||:||:||||:|  | :|  |||: ||| :   ||  |:
Db         360 ISRLLKHNPSQRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 33
US-10-087-684-50
; Sequence 50, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 50
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-50

  Query Match            58.1%;  Score 1050.5;  DB 15;  Length 403;
  Best Local Similarity   61.7%;  Pred. No. 1.2e-77;
  Matches  209;  Conservative   40;  Mismatches   67;  Indels   23;  Gaps    5;

Qy           4 KENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSSG 63
               :: | | |  |          |      |:|: |||      :| :   |  ||  |:
Db          80 QQTSVPHPVSR--------PLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK- 125

Qy          64 TPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
                 | : ::|||||||||||||||||||:|  |:|||||:|:|| |||||||
Db         126 ------RQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLR 179

Qy         124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
               ||:|||:|| |||||||| ||:|  |:||||||| | :|:|||   |||||||  : |
Db         180 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITE 239

Qy         184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
```

```
              ||:|| ||| |:|||||||||||||||  ||||||||||||||| || |:||||||||||
Db        240 LANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPE 299

Qy        244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
              ||||||:|||||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||
Db        300 MIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDL 359

Qy        304 ISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              ||:||:||||:|  | :|  |||: ||| :   ||  |:
Db        360 ISRLLKHNPSQRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 34
US-10-218-779-50
; Sequence 50, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-50

  Query Match             58.1%;  Score 1050.5;  DB 15;  Length 403;
  Best Local Similarity   61.7%;  Pred. No. 1.2e-77;
  Matches  209;  Conservative   40;  Mismatches   67;  Indels   23;  Gaps    5;

Qy          4 KENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSSG 63
              :: | | |  |          |      |:|: |||     :| :   |  ||  |:
Db         80 QQTSVPHPVSR--------PLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK- 125

Qy         64 TPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
                 | : ::|||||||||||||||||||||:|  |:|||||||:|:|| |||||||
Db        126 ------RQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLR 179

Qy        124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
              ||:|||:|| |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |
Db        180 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITE 239

Qy        184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
              ||:|| ||| |:|||||||||||||  |||||||||||||| || |:||||||||||
Db        240 LANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPE 299

Qy        244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
              |||||||:|||||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||
Db        300 MIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDL 359

Qy        304 ISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              ||:||:||||:|  | :|  |||: ||| :   ||  |:
Db        360 ISRLLKHNPSQRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 35
US-10-087-684-51
; Sequence 51, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 51
;   LENGTH: 402
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

```
US-10-087-684-51

  Query Match             57.8%;  Score 1045.5;  DB 15;  Length 402;
  Best Local Similarity   54.8%;  Pred. No. 3.2e-77;
  Matches  216;  Conservative   44;  Mismatches   75;  Indels   59;  Gaps    8;

Qy          4 KENSYPWPYGRQTAPSG-------------LSTLP------QRVL------RKEPVTPSA 38
              |||     |  : ||| |               : ||      ||||      :: |:
Db          5 KENCISGPV-KATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQAQK 63

Qy         39 LV-------------LMSRSNVQPTAAPGQKVMENSSGTPD----ILTRHF--------- 72
              ||              | : |   | : |     ::    |     || |::
Db         64 LVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSHLKIILRRNWHQNRKMKNQ 123

Qy         73 ----TIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEI 128
                  ::|||||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:||
Db        124 KEAVALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEI 183

Qy        129 QAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADAL 188
              |:|| |||||||| ||:|  |:||||||| | :|:||||   ||||||| :   :|:||
Db        184 QSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTANLYNRIANAL 243

Qy        189 MYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGR 248
              ||| |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||
Db        244 SYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGR 303

Qy        249 MHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLL 308
              ||:|||||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||
Db        304 MHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLL 363

Qy        309 RHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :||||:|  | :|  |||: ||| :   ||  |:
Db        364 KHNPSQRPMLREVLEHPWITANSSK---PSNCQN 394


RESULT 36
US-10-218-779-51
; Sequence 51, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
;   LENGTH: 402
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-51

  Query Match             57.8%;  Score 1045.5;  DB 15;  Length 402;
  Best Local Similarity   54.8%;  Pred. No. 3.2e-77;
  Matches  216;  Conservative   44;  Mismatches   75;  Indels   59;  Gaps    8;

Qy          4 KENSYPWPYGRQTAPSG-------------LSTLP------QRVL------RKEPVTPSA 38
              |||     |  : ||| |               : ||      ||||      :: |:
Db          5 KENCISGPV-KATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQAQK 63

Qy         39 LV-------------LMSRSNVQPTAAPGQKVMENSSGTPD----ILTRHF--------- 72
              ||              | : |   | : |     ::    |     || |::
Db         64 LVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSHLKIILRRNWHQNRKMKNQ 123

Qy         73 ----TIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEI 128
                  ::|||||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:||
Db        124 KEAVALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEI 183

Qy        129 QAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADAL 188
              |:|| |||||||| ||:|  |:||||||| | :|:||||   ||||||| :   :|:||
Db        184 QSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTANLYNRIANAL 243

Qy        189 MYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGR 248
              ||| |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||
Db        244 SYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGR 303

Qy        249 MHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLL 308
              ||:|||||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||
Db        304 MHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLL 363

Qy        309 RHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :||||:|  | :|  |||: ||| :   ||  |:
Db        364 KHNPSQRPMLREVLEHPWITANSSK---PSNCQN 394


RESULT 37
US-10-087-684-14
; Sequence 14, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R.
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Grosse, William M.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Cathereine E.
;  APPLICANT:  Casman, Stacie, J.
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Boldog, Ferenc L.
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
```

```
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Gangolli, Esha A.
;  TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-214 CIP
;  CURRENT APPLICATION NUMBER: US/10/087,684
;  CURRENT FILING DATE:  2003-03-10
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/274,194
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 220
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 14
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-087-684-14

  Query Match             53.4%;  Score 966;  DB 15;  Length 403;
  Best Local Similarity   61.8%;  Pred. No. 1e-70;
  Matches  186;  Conservative   41;  Mismatches   68;  Indels    6;  Gaps    2;

Qy          45 SNVQPTAAP---GQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHF 101
               |   | :||    :: : :     :   | : ::| ||||| ||||||||||||||:| |
Db          98 SKQSPLSAPENNPEEELASKQKNEESKKRQWALEDLEIGRPPGKGKFGNVYLAREKQSKF 157

Qy         102 IVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGE 161
               |:||:||||:|:|| |||||||||:|||:|| ||||:||| ||:|  |:|||||
Db         158 ILALRVLFKAQLEKAGVEHQLRREVEIQSHLQHPNIIRLYGYFHDATRVYLILEYTPLET 217

Qy         162 LYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGW 221
               :  ||||   |||||||| : ||| || ||| | ||||||||||||||  |||:||:|||
Db         218 VNTELQKLSKFDEQRTATYITELASALSYCHSKTVIHRDIKPENLLLGSAGELEIANFGW 277

Qy         222 SVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNET 281
               | |||| || |:|||||||||||||||||:|||||| :|||| | ||| |||: :: ||
Db         278 SEHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCCEFLVGKPPFEANTYQET 337

Qy         282 YRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQ 341
               |:|| :|:  ||  |  ||:||||:||:| ||:|  | :|  :||: ||| :   ||  |
Db         338 YKRISRVEFTFPDFVTEGARDLISRLLKHVPSQRPMLREVLEYPWITANSSK---PSNCQ 394

Qy         342 S 342
               :
Db         395 N 395


RESULT 38
US-10-218-779-14
; Sequence 14, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Stone, David
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William
;  APPLICANT:  Alsobrook II, John
;  APPLICANT:  Lepley, Denise
;  APPLICANT:  Rieger, Daniel
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Casman, Stacie
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Li, Li
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shenoy, Suresh
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Zerhusen, Bryan
;  APPLICANT:  Malyankar, Uriel
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Gangolli, Esha
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-214
;  CURRENT APPLICATION NUMBER: US/10/218,779
;  CURRENT FILING DATE:  2002-08-14
;  PRIOR APPLICATION NUMBER: 60/253,834
;  PRIOR FILING DATE: 2000-11-29
;  PRIOR APPLICATION NUMBER: 60/250,-926
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: 60/264,180
;  PRIOR FILING DATE: 2001-01-25
;  PRIOR APPLICATION NUMBER: 60/313,656
;  PRIOR FILING DATE: 2001-08-20
;  PRIOR APPLICATION NUMBER: 60/327,456
;  PRIOR FILING DATE: 2001-10-05
;  NUMBER OF SEQ ID NOS: 216
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 403
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-218-779-14

  Query Match             53.4%;  Score 966;  DB 15;  Length 403;
  Best Local Similarity   61.8%;  Pred. No. 1e-70;
  Matches  186;  Conservative   41;  Mismatches   68;  Indels    6;  Gaps    2;

Qy          45 SNVQPTAAP---GQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHF 101
               |   | :||    :: : :     :   | : ::| ||||| ||||||||||||||:| |
Db          98 SKQSPLSAPENNPEEELASKQKNEESKKRQWALEDLEIGRPPGKGKFGNVYLAREKQSKF 157

Qy         102 IVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGE 161
               |:||:||||:|:|| |||||||||:|||:|| ||||:||| ||:|  |:|||||
Db         158 ILALRVLFKAQLEKAGVEHQLRREVEIQSHLQHPNIIRLYGYFHDATRVYLILEYTPLET 217

Qy         162 LYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGW 221
               :  ||||   |||||||| : ||| || ||| | ||||||||||||||  |||:||:|||
Db         218 VNTELQKLSKFDEQRTATYITELASALSYCHSKTVIHRDIKPENLLLGSAGELEIANFGW 277

Qy         222 SVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNET 281
               | |||| || |:|||||||||||||||||:|||||| :|||| | ||| |||: :: ||
Db         278 SEHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCCEFLVGKPPFEANTYQET 337

Qy         282 YRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQ 341
               |:|| :|:  ||  |  ||:||||:||:| ||:|  | :|  :||: ||| :   ||  |
Db         338 YKRISRVEFTFPDFVTEGARDLISRLLKHVPSQRPMLREVLEYPWITANSSK---PSNCQ 394

Qy         342 S 342
               :
Db         395 N 395


RESULT 39
US-10-389-566-1246
; Sequence 1246, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1246
;   LENGTH: 294
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-389-566-1246

  Query Match           52.0%;  Score 940;  DB 15;  Length 294;
  Best Local Similarity   60.3%;  Pred. No. 9.7e-69;
  Matches  179;  Conservative   47;  Mismatches   63;  Indels    8;  Gaps    4;

Qy         49 PTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVL 108
              ||    |:  | |  :     : :|: ||:||:|||:||||:|||||||:|: :||||||
Db          4 PTETQHQE-KEASDASAAAAQKRWTLSDFDIGKPLGRGKFGHVYLAREKRSNHVVALKVL 62

Qy        109 FKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQK 168
              ||||:::  |||||||:|||:|| |||||||| ||||::|:||||||| ||||||:|||
Db         63 FKSQLQQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQKRVYLILEYAARGELYKDLQK 122

Qy        169 SCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSL 228
                | |:| || :  || ||:||||| ||||||||||||:| :|||||||||||||  :
Db        123 CKYFSERRAATYVASLARALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN- 181

Qy        229 RRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKV 288
              ||:|||||||||||:|   |:  ||:| :|:|||| | | ||||:  |::||||||:|
Db        182 RRRTMCGTLDYLPPEMVESVEHDASVDIWSLGILCYEFLYGVPPFEAMEHSDTYRRIVQV 241

Qy        289 DLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
              |||||    :   |:||||::|    |:|||| ::  |||:  |:     || :  |
Db        242 DLKFPPKPIISASAKDLISQMLVKESSQRLPLHKLLEHPWIVQNA----DPSGIYRV 294


RESULT 40
US-10-389-566-1606
; Sequence 1606, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1606
;   LENGTH: 294
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (182)..(182)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-1606

  Query Match           51.8%;  Score 937;  DB 15;  Length 294;
  Best Local Similarity   60.3%;  Pred. No. 1.7e-68;
  Matches  179;  Conservative   46;  Mismatches   64;  Indels    8;  Gaps    4;

Qy         49 PTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVL 108
              ||    |:  | |  :     : :|: ||:||:|||:||||:|||||||:|: :||||||
Db          4 PTETQHQE-KEASDASAAAAQKRWTLSDFDIGKPLGRGKFGHVYLAREKRSNHVVALKVL 62

Qy        109 FKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQK 168
              ||||:::  |||||||:|||:|| |||||||| ||||::|:||||||| ||||||:|||
Db         63 FKSQLQQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQKRVYLILEYAARGELYKDLQK 122

Qy        169 SCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSL 228
                | |:| || :  || ||:||||| ||||||||||||:| :|||||||||||||  :
Db        123 CKYFSERRAATYVASLARALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFNX 182

Qy        229 RRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKV 288
              || |||||||||||||:|   |:  ||:| :|:|||| | | ||||:  |::||||||:|
Db        183 RR-TMCGTLDYLPPEMVESVEHDASVDIWSLGILCYEFLYGVPPFEAMEHSDTYRRIVQV 241

Qy        289 DLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
              |||||    :   |:||||::|    |:|||| ::  |||:  |:     || :  |
Db        242 DLKFPPKPIISASAKDLISQMLVKESSQRLPLHKLLEHPWIVQNA----DPSGIYRV 294


RESULT 41
US-10-389-566-645
; Sequence 645, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 645
;   LENGTH: 295
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-389-566-645

  Query Match           51.7%;  Score 934.5;  DB 15;  Length 295;
  Best Local Similarity   65.5%;  Pred. No. 2.7e-68;
  Matches  173;  Conservative   44;  Mismatches   44;  Indels    3;  Gaps    2;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              | :|::||:||:|||:||||:|||||||  |: ||||||||||||:::  | ||||||:|||
Db         25 RRWTLNDFDIGKPLGRGKFGHVYLAREKTSNHIVALKVLFKSQLQQSQVVHQLRREVEIQ 84

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| ||:||||| ||||::|:|||||||:|||||||    | |:| || :  || ||:
Db         85 SHLRHPHILRLYGYFYDQKRVYLILEYAPKGELYKELQKCKYFSERRAATYVASLARALI 144

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        145 YCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 203

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|||||  | | ||||:  |::|||||::||||||    | : |:||||::
Db        204 HDASVDIWSLGVLCYEFLYGVPPFEAKEHSDTYRRIIQVDLKFPPKPIVSSAAKDLISQM 263

Qy        308 LRHNPSERLPLAQVSAHPWVRANS 331
              |  : |:|||| ::  |||:  |:
```

```
Db        264 LVKDSSQRLPLHKLLEHPWIVQNA 287


RESULT 42
US-10-389-566-644
; Sequence 644, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 644
;   LENGTH: 298
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-389-566-644

  Query Match             51.6%;  Score 933.5;  DB 15;  Length 298;
  Best Local Similarity   65.5%;  Pred. No. 3.3e-68;
  Matches  173;  Conservative   44;  Mismatches   44;  Indels    3;  Gaps    2;

Qy          70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
               | :|::||:||:|||:||||:||||||| |: ||||||||||:::  | ||||||:|||
Db          28 RRWTLNDFDIGKPLGRGKFGHVYLAREKTSNHIVALKVLFKSQLQQSQVVHQLRREVEIQ 87

Qy         130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
               :|| ||:|||| ||||::|:|||||||:|||||||    | |:| || :  || ||:
Db          88 SHLRHPHILRLYGYFYDQKRVYLILEYAPKGELYKELQKCKYFSERRAATYVASLARALI 147

Qy         190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
               ||||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db         148 YCHGKHVIHRDIKPENLLIGSQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 206

Qy         250 HNEKVDLWCIGVLCYELLVGNPPPESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
               |:  ||:| :|||||| | | |||||:  |::|||||::||||||    | : |:||||::
Db         207 HDASVDIWSLGVLCYEFLYGVPPPEAKEHSDTYRRIIQVDLKFPPKPIVSSAAKDLISQM 266

Qy         308 LRHNPSERLPLAQVSAHPWVRANS 331
               |  : |:|||| ::  |||:  |:
Db         267 LVKDSSQRLPLHKLLEHPWIVQNA 290


RESULT 43
US-10-389-566-1240
; Sequence 1240, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1240
;   LENGTH: 282
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-389-566-1240

  Query Match             50.9%;  Score 921.5;  DB 15;  Length 282;
  Best Local Similarity   63.7%;  Pred. No. 3e-67;
  Matches  174;  Conservative   42;  Mismatches   50;  Indels    7;  Gaps    3;

Qy          70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
               : :|  ||:||:|||:||||:|||||||:|  |||||||||:|:::  |||||||:|||
Db          12 KRWTTSDFDIGKPLGRGKFGHVYLAREKRSDHIVALKVLFKAQLQQSQVEHQLRREVEIQ 71

Qy         130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
               :|| |||||||| ||||::|:|||||| ||||||||||    | |:| || :  || ||:
Db          72 SHLRHPNILRLYGYFYDQKRVYLILEYAVRGELYKELQKCKYFSERRAATYVASLARALI 131

Qy         190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
               ||||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db         132 YCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 190

Qy         250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
               |:  ||:| :|:|||| | | ||||:  |:|||:|||:||||||    | : |:||||::
Db         191 HDASVDIWSLGILCYEFLYGVPPFEAREHSETYKRIVQVDLKFPPKPIVSSSAKDLISQM 250

Qy         308 LRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
               |    ::|| | ::  |||:  |:     || |
Db         251 LVKESTQRLALHKLLEHPWIVQNA----DPSGL 279


RESULT 44
US-10-389-566-382
; Sequence 382, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 382
;   LENGTH: 368
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (48)..(48)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (56)..(56)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (86)..(87)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
```

US-10-389-566-382

Query Match 50.7%; Score 918; DB 15; Length 368;
Best Local Similarity 58.8%; Pred. No. 8e-67;
Matches 177; Conservative 47; Mismatches 59; Indels 18; Gaps 4;

```
Qy         48 QPT--AAPGQKVMENSSGTPDIL----------TRHFTIDDFEIGRPLGKGKFGNVY 92
              :||  ::|   :: |  |   :             : : : |||:|:|||:||||:||
Db         57 RPTHASSPSAPLLHNHGGRHRVARQRGQGXXSHANQEKRWVLSDFEVGKPLGRGKFGHVY 116

Qy         93 LAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYL 152
              |||||:|  |||||||||||:::  ||||||||:|||:|| |||||||| ||||: |:||
Db        117 LAREKRSSEIVALKVLFKSQLKQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQTRVYL 176

Qy        153 ILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKG 212
              ||||| :|||||||  :   | |:|||| :  || ||:| ||| ||||||||||||:|::|
Db        177 ILEYAAKGELYKELTRCKHFSERRTATYVASLARALIYLHGKHVIHRDIKPENLLVGVQG 236

Qy        213 ELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPP 272
              |:||||||||||  : ||:|||||||||||||:|   |:  ||:| :|||||| | | ||
Db        237 EIKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVEKAEHDYHVDIWSLGVLCYEFLYGVPP 295

Qy        273 FESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRAN 330
              ||:  |:|||||||||||||    |   |:||||::|  : ::|||| :|  |||:  |
Db        296 FEAKEHSETYRRIVKVDLKFPLKPFVSPAAKDLISQMLVKSSAQRLPLHKVLEHPWIVQN 355

Qy        331 S 331
              :
Db        356 A 356
```

RESULT 45
US-10-389-566-819
; Sequence 819, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 819
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (144)..(144)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-819

Query Match 50.7%; Score 917.5; DB 15; Length 418;
Best Local Similarity 65.2%; Pred. No. 1e-66;
Matches 172; Conservative 41; Mismatches 48; Indels 3; Gaps 2;

```
Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              : : : ||:||:|||:||||:|||||||:|: ||||||||||:::  ||||||||:|||
Db        148 KRWVLSDFDIGKPLGRGKFGHVYLAREKRSNHIVALKVLFKSQLKQSQVEHQLRREVEIQ 207

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| |||||||| |||| |:|||||||| :||||||||:    | |:|:|| :  || ||:
Db        208 SHLRHPNILRLYGYFYDQTRVYLILEYALKGELYKELQRCKHFSERRSATYIASLAHALI 267

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              | ||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        268 YLHGKHVIHRDIKPENLLIGSQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVEKTE 326

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|:|||| | | ||||:  |:||||||||||||     |   |:||||::
Db        327 HDYHVDIWSLGILCYEFLYGVPPFEAKEHSETYRRIVKVDLKFPLKPFVSPAAKDLISQM 386

Qy        308 LRHNPSERLPLAQVSAHPWVRANS 331
              |  | : |||| ::  |||:  |:
Db        387 LVKNSAHRLPLHKLLEHPWIVQNA 410
```

RESULT 46
US-10-437-963-133574
; Sequence 133574, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133574
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(320)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35430C.1.pep
US-10-437-963-133574

Query Match 50.7%; Score 916.5; DB 16; Length 320;
Best Local Similarity 65.2%; Pred. No. 9e-67;
Matches 172; Conservative 41; Mismatches 48; Indels 3; Gaps 2;

```
Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              : : : ||:||:|||:||||:|||||||:|: ||||||||||:::  ||||||||:|||
Db         50 KRWVLSDFDIGKPLGRGKFGHVYLAREKRSNHIVALKVLFKSQLKQSQVEHQLRREVEIQ 109

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| |||||||| |||| |:|||||||| :||||||||:    | |:|:|| :  || ||:
Db        110 SHLRHPNILRLYGYFYDQTRVYLILEYALKGELYKELQRCKHFSERRSATYIASLARALI 169

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              | ||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        170 YLHGKHVIHRDIKPENLLIGSQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVEKTE 228

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|:|||| | | ||||:  |:||||||||||||     |   |:||||::
Db        229 HDYHVDIWSLGILCYEFLYGVPPFEAKEHSETYRRIVKVDLKFPLKPFVSPAAKDLISQM 288

Qy        308 LRHNPSERLPLAQVSAHPWVRANS 331
              |  | : |||| ::  |||:  |:
Db        289 LVKNSAHRLPLHKLLEHPWIVQNA 312
```

RESULT 47
US-10-424-599-274736
; Sequence 274736, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

```
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 274736
;   LENGTH: 307
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(307)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_90107C.1.pep
US-10-424-599-274736

  Query Match             50.2%;  Score 909;  DB 15;  Length 307;
  Best Local Similarity   62.9%;  Pred. No. 3.5e-66;
  Matches  173;  Conservative   44;  Mismatches   44;  Indels   14;  Gaps    4;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              | :|::||:||:|||:||||:||||||| |: |||||||||||::  | ||||||:|||
Db         26 RRWTLNDFDIGKPLGRGKFGHVYLAREKTSNHIVALKVLFKSQLQQSQVVHQLRREVEIQ 85

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEEL----- 184
              :|| ||:||||| ||||::|:||||||||:|||||||||   | |:| || :  |
Db         86 SHLRHPHILRLYGYFYDQKRVYLILEYAPKGELYKELQKCKYFSERRAATYVASLELHLR 145

Qy        185 -ADALMYCHGKK-----VIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLD 238
               | ||:|||||      |||||||||||:| :|||||||||||||  : ||:|||||||
Db        146 FARALIYCHGKXEVIHFVIHRDIKPENLLMGAEGELKIADFGWSVHTFN-RRRTMCGTLD 204

Qy        239 YLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--V 296
              |||||||:|   |:  ||:| :|||||| | | ||||:  |::|||||::||||||    |
Db        205 YLPPEMVESVEHDASVDIWSLGVLCYEFLYGVPPFEAKEHSDTYRRIIQVDLKFPPKPIV 264

Qy        297 PTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANS 331
               : |:||||::|  : |:|||| ::  |||:  |:
Db        265 SSAAKDLISQMLVKDSSQRLPLHKLLEHPWIVQNA 299


RESULT 48
US-10-389-566-550
; Sequence 550, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 550
;   LENGTH: 298
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (130)..(130)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (135)..(135)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-550

  Query Match             50.2%;  Score 908;  DB 15;  Length 298;
  Best Local Similarity   63.7%;  Pred. No. 4.1e-66;
  Matches  170;  Conservative   45;  Mismatches   46;  Indels    6;  Gaps    3;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              | :|::||:||:|||:||||:||||||| |: |||||||||||::  | ||||||:|||
Db         25 RRWTLNDFDIGKPLGRGKFGHVYLAREKTSNHIVALKVLFKSQLQQSQVVHQLRREVEIQ 84

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTF---DEQRTATIMEELAD 186
              :|| ||:| ||| ||||::|:||||||||:||||||||   |   : :  || :  ||
Db         85 SHLRHPHIWRLYGYFYDQKRVYLILEYAPKGELYKELQKCKYFSRXERRAXATYVASLAR 144

Qy        187 ALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIE 246
              ||:||||| |||||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        145 ALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVE 203

Qy        247 GRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLI 304
                 |:  ||:| :|||||| | | ||||:  |::|||||::||||||    | : |:|||
Db        204 SVEHDASVDIWSLGVLCYEFLYGVPPFEAKEHSDTYRRIIQVDLKFPPKPIVSSAAKDLI 263

Qy        305 SKLLRHNPSERLPLAQVSAHPWVRANS 331
              |::|  : |:|||| ::  |||:  |:
Db        264 SQMLVKDSSQRLPLHKLLEHPWIVQNA 290


RESULT 49
US-10-389-566-2420
; Sequence 2420, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2420
;   LENGTH: 432
;   TYPE: PRT
;   ORGANISM: Oryza sativa
US-10-389-566-2420

  Query Match             50.1%;  Score 906;  DB 15;  Length 432;
  Best Local Similarity   65.2%;  Pred. No. 9.3e-66;
  Matches  172;  Conservative   40;  Mismatches   48;  Indels    4;  Gaps    3;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              : : : ||:||:|||:||||:||||||:|: |||||||||||::  |||||||:|||
Db        163 KRWVLSDFDIGKPLGRGKFGHVYLAREKRSNHIVALKVLFKSQLKQSQVEHQLRREVEIQ 222

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| |||||||| |||| |:|||||| :||||||||:    | |:|:|| :  || ||:
Db        223 SHLRHPNILRLYGYFYD-TRVYLILEYALKGELYKELQRCKHFSERRSATYIASLAHALI 281

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              | ||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        282 YLHGKHVIHRDIKPENLLIGSQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVEKTE 340
```

```
Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|:|||| | | ||||:  |:||||||||||||||    |   |:||||::
Db        341 HDYHVDIWSLGILCYEFLYGVPPFEAKEHSETYRRIVKVDLKFPLKPFVSPAAKDLISQM 400

Qy        308 LRHNPSERLPLAQVSAHPWVRANS 331
              |  | : |||| ::  |||:  |:
Db        401 LVKNSAHRLPLHKLLEHPWIVQNA 424
```

Search completed: June 20, 2005, 20:06:50
Job time : 72.8132 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:57:12 ; Search time 118.351 Seconds
(without alignments)
1488.417 Million cell updates/sec

Title: US-10-734-126-3
Perfect score: 1809
Sequence: 1 MAQKENSYPWPYGRQTAPSG.....PWVRANSRRVLPPSALQSVA 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 33

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1809	100.0	344	1	AURB_HUMAN	Q96gd4 homo sapien
2	1670	92.3	344	2	Q7YRC6	Q7yrc6 bos taurus
3	1664	92.0	344	2	Q7YRC7	Q7yrc7 sus scrofa
4	1525.5	84.3	345	1	AURB_MOUSE	O70126 mus musculu
5	1520.5	84.1	343	1	AURB_RAT	O55099 rattus norv
6	1519.5	84.0	345	2	Q8C6C1	Q8c6c1 mus musculu
7	1239	68.5	361	2	Q6DE08	Q6de08 xenopus lae
8	1239	68.5	371	2	Q7ZYT9	Q7zyt9 xenopus lae
9	1238	68.4	361	2	Q9DF70	Q9df70 xenopus lae
10	1237	68.4	361	2	Q8JG74	Q8jg74 xenopus lae
11	1234.5	68.2	368	2	Q6GPL3	Q6gpl3 xenopus lae
12	1230	68.0	309	1	AURC_HUMAN	Q9uqb9 homo sapien
13	1222	67.6	289	2	Q6AZY8	Q6azy8 homo sapien
14	1222	67.6	290	2	Q6DLZ0	Q6dlz0 homo sapien
15	1156	63.9	320	2	Q6NW76	Q6nw76 brachydanio
16	1154	63.8	282	1	AURC_MOUSE	O88445 mus musculu
17	1147	63.4	276	2	Q6P209	Q6p209 mus musculu
18	1111.5	61.4	346	2	Q8JGS8	Q8jgs8 brachydanio
19	1084	59.9	395	1	STK6_MOUSE	P97477 mus musculu
20	1084	59.9	395	2	Q7TNK2	Q7tnk2 mus musculu
21	1080	59.7	386	2	Q8C3H8	Q8c3h8 mus musculu
22	1073	59.3	395	2	Q8BP87	Q8bp87 mus musculu
23	1071	59.2	397	1	STK6_RAT	P59241 rattus norv
24	1062	58.7	408	1	ST6L_XENLA	Q91819 xenopus lae
25	1060	58.6	408	2	Q6DJK0	Q6djk0 xenopus lae
26	1057.5	58.5	403	1	STK6_HUMAN	O14965 homo sapien
27	1050.5	58.1	407	1	STK6_XENLA	Q91820 xenopus lae
28	1041.5	57.6	405	2	Q6DBZ4	Q6dbz4 brachydanio
29	940	52.0	294	2	Q9M077	Q9m077 arabidopsis
30	937	51.8	294	2	Q8LBX4	Q8lbx4 arabidopsis
31	921.5	50.9	282	2	O82309	O82309 arabidopsis
32	921.5	50.9	288	2	Q683C9	Q683c9 arabidopsis
33	906	50.1	432	2	Q9LG71	Q9lg71 oryza sativ

ALIGNMENTS

RESULT 1
AURB_HUMAN
ID AURB_HUMAN STANDARD; PRT; 344 AA.
AC Q96GD4; O14630; O60446; O95083; Q96DV5; Q9UQ46;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase 12 (EC 2.7.1.37) (Aurora- and Ipl1-
DE like midbody-associated protein 1) (AIM-1) (Aurora/IPL1-related kinase
DE 2) (Aurora-related kinase 2) (STK-1) (Aurora-B).
GN Name=AURKB; Synonyms=AIK2, AIM1, ARK2, STK12;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98183439; PubMed=9514916; DOI=10.1006/bbrc.1998.8250;
RA Shindo M., Nakano H., Kuroyanagi H., Shirasawa T., Mihara M.,
RA Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;
RT "cDNA cloning, expression, subcellular localization, and chromosomal
RT assignment of mammalian aurora homologues, aurora-related kinase (ARK)
RT 1 and 2.";
RL Biochem. Biophys. Res. Commun. 244:285-292(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99025855; PubMed=9809983;
RA Tatsuka M., Katayama H., Ota T., Tanaka T., Odashima S., Suzuki F.,
RA Terada Y.;
RT "Multinuclearity and increased ploidy caused by overexpression of the
RT aurora- and Ipl1-like midbody-associated protein mitotic kinase in
RT human cancer cells.";
RL Cancer Res. 58:4811-4816(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Spleen;
RX MEDLINE=99077743; PubMed=9858806;
RA Kimura M., Matsuda Y., Yoshioka T., Sumi N., Okano Y.;
RT "Identification and characterization of STK12/Aik2: a human gene
RT related to aurora of Drosophila and yeast IPL1.";
RL Cytogenet. Cell Genet. 82:147-152(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=21364697; PubMed=11471245;
RA Prigent C., Gill R., Trower M., Sanseau P.;
RT "In silico cloning of a new protein kinase, Aik2, related to
RT Drosophila aurora using the new tool: EST Blast.";
RL In Silico Biol. 1:123-128(1999).
RN [5]
RP SEQUENCE FROM N.A.
RA Zhang Q., Yu L., Bi A.;
RT "Cloning of a novel human gene homologous to mouse STK-1.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung, Lymph, and Muscle;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   REVIEW.
RX   MEDLINE=21306577; PubMed=11413462; DOI=10.1038/35048096;
RA   Nigg E.A.;
RT   "Mitotic kinases as regulators of cell division and its checkpoints.";
RL   Nat. Rev. Mol. Cell Biol. 2:21-32(2001).
CC   -!- FUNCTION: May be directly involved in regulating the cleavage of
CC       polar spindle microtubules and is a key regulator for the onset of
CC       cytokinesis during mitosis.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localized to the midzone of central spindle
CC       in late anaphase and concentrated into the midbody in telophase
CC       and cytokinesis. Colocalized with gamma tubulin in the mid-body.
CC   -!- TISSUE SPECIFICITY: High level expression seen in the thymus. It
CC       is also expressed in the spleen, lung, testis, colon, placenta and
CC       fetal liver. Expressed during S and G2/M phase and expression is
CC       up-regulated in cancer cells during M phase.
CC   -!- DISEASE: Disruptive regulation of expression is a possibile
CC       mechanism of the perturbation of chromosomal integrity in cancer
CC       cells through its dominant-negative effect on cytokinesis.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF008552; AAC12709.1; -.
DR   EMBL; AB011450; BAA32136.1; -.
DR   EMBL; AB011446; BAA82709.1; -.
DR   EMBL; AF004022; AAB65786.1; -.
DR   EMBL; AF015254; AAC98891.1; -.
DR   EMBL; BC000442; AAH00442.2; ALT_INIT.
DR   EMBL; BC009751; AAH09751.1; -.
DR   EMBL; BC013300; AAH13300.2; ALT_INIT.
DR   HSSP; P31751; 1GZK.
DR   Genew; HGNC:11390; AURKB.
DR   H-InvDB; HIX0019005; -.
DR   MIM; 604970; -.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       77    327       Protein kinase.
FT   NP_BIND      83     91       ATP (By similarity).
FT   BINDING     106    106       ATP (By similarity).
FT   ACT_SITE    200    200       Proton acceptor (By similarity).
FT   CONFLICT     14     15       RQ -> DK (in Ref. 5).
FT   CONFLICT     70     70       R -> RR (in Ref. 6; AAH13300).
FT   CONFLICT    161    161       E -> M (in Ref. 4 and 5).
FT   CONFLICT    167    169       QKS -> HKT (in Ref. 4).
FT   CONFLICT    179    179       T -> TVRR (in Ref. 4).
FT   CONFLICT    180    180       I -> VRAV (in Ref. 5).
FT   CONFLICT    226    226       P -> T (in Ref. 3).
FT   CONFLICT    249    250       MH -> ID (in Ref. 3).
FT   CONFLICT    271    271       Missing (in Ref. 3).
FT   CONFLICT    298    298       T -> M (in Ref. 6; AAH09751/AAH13300).
SQ   SEQUENCE   344 AA;  39280 MW;  8325E3EF5A1FB170 CRC64;

  Query Match             100.0%;  Score 1809;  DB 1;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-123;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 2
Q7YRC6
ID   Q7YRC6      PRELIMINARY;      PRT;   344 AA.
AC   Q7YRC6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Serine/threonine kinase 12.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G., Li W., Yu L.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY336975; AAQ16151.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
```

```
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   344 AA;  39441 MW;  4DD7158CF2F5D047 CRC64;

  Query Match             92.3%;  Score 1670;  DB 2;  Length 344;
  Best Local Similarity   93.0%;  Pred. No. 2.8e-113;
  Matches  319;  Conservative    7;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||:|||||||||||  ||:|||||||||||||||||||||| ||||||||||:||
Db          1 MAQKENAYPWPYGRQTAQPGLNTLPQRVLRKEPVTPSALVLMSRSNAQPTAAPGQKVVEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              |||||:|  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPNIPKRSFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              |||||||||||| ||||||||||||||||||||||||||||||| ||||||||||
Db        121 QLRREIEIQAHLQHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSRTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              |||||||| ||| ||||||||||||||:||||||||||||||||||||||||||||||
Db        181 MEELADALTYCHAKKVIHRDIKPENLLLGLRGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              |||||||| |||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        241 PPEMIEGRTHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPPSVPLGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
              || | |||:||||||||||||||||| :||||||||| |||
Db        301 QDFIYKLLKHNPSERLPLAQVSAHPWVRTHSRRVLPPSAPQSV 343


RESULT 3
Q7YRC7
ID   Q7YRC7      PRELIMINARY;      PRT;   344 AA.
AC   Q7YRC7;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Serine/threonine kinase 12.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G., Li W., Yu L.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY336974; AAQ16150.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   344 AA;  39275 MW;  5CEC29C50C307F12 CRC64;

  Query Match             92.0%;  Score 1664;  DB 2;  Length 344;
  Best Local Similarity   92.7%;  Pred. No. 7.6e-113;
  Matches  318;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||:|||||||||||  |||: |||||||| |||||||||||||| ||||||||||:||
Db          1 MAQKENTYPWPYGRQTAQSGLNILPQRVLRKEAVTPSALVLMSRSNTQPTAAPGQKVVEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              |||||:  || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPNFSTRSFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              |||||||||||| |||||||||||||||||||||||||||||| ||||||||  ||||||||||
Db        121 QLRREIEIQAHLQHPNILRLYNYFYDRRRIYLILEYAPRXELYKELQKCRTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||:|||||||||||||||||||:|||||||||||||||||||||||||||||||
Db        181 MEELADALIYCHGKKVIHRDIKPENLLLGLQGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPPSVPTGA 300
              |||||||| |||||||||||||||||||||||||||||||||||||||||| ||| ||
Db        241 PPEMIEGRTHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIGKVDLKFPPSVPAGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
              |||||||:||||:|||||||||||||:||||||||| |||
Db        301 QDLISKLLKHNPSDRLPLAQVSAHPWVRAHSRRVLPPSAPQSV 343


RESULT 4
AURB_MOUSE
ID   AURB_MOUSE      STANDARD;      PRT;   345 AA.
AC   O70126; Q61882;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase 12 (EC 2.7.1.37) (Aurora-related
DE   kinase 2) (Serine/threonine-protein kinase 5) (STK-1) (Aurora-B).
GN   Name=Aurkb; Synonyms=Ark2, Stk1, Stk12, Stk5;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Testis;
RX   MEDLINE=96194801; PubMed=8647446; DOI=10.1016/0378-1119(95)00809-8;
RA   Niwa H., Abe K., Kunisada T., Yamamura K.;
RT   "Cell-cycle-dependent expression of the STK-1 gene encoding a novel
RT   murine putative protein kinase.";
RL   Gene 169:197-201(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=98183439; PubMed=9514916; DOI=10.1006/bbrc.1998.8250;
RA   Shindo M., Nakano H., Kuroyanagi H., Shirasawa T., Mihara M.,
RA   Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;
RT   "cDNA cloning, expression, subcellular localization, and chromosomal
RT   assignment of mammalian aurora homologues, aurora-related kinase (ARK)
RT   1 and 2.";
RL   Biochem. Biophys. Res. Commun. 244:285-292(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
```

```
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May be directly involved in regulating the cleavage of
CC       polar spindle microtubules and is a key regulator for the onset of
CC       cytokinesis during mitosis (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localized to the midzone of central spindle
CC       in late anaphase and concentrated into the midbody in telophase
CC       and cytokinesis. Colocalized with gamma tubulin in the mid-body
CC       (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in testis, intestine and spleen. All
CC       of them are tissues that contain a large number of proliferating
CC       cells. Expressed during S phase, in a cell-cycle-dependent
CC       fashion.
CC   -!- DEVELOPMENTAL STAGE: Strongly expressed in 8.5 and 12.5 dpc.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D21099; BAA04658.1; -.
DR   EMBL; U69107; AAC12683.1; -.
DR   EMBL; BC003261; AAH03261.1; -.
DR   HSSP; P31751; 1GZK.
DR   MGD; MGI:107168; Aurkb.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       82    332       Protein kinase.
FT   NP_BIND      88     96       ATP (By similarity).
FT   BINDING     111    111       ATP (By similarity).
FT   ACT_SITE    205    205       Proton acceptor (By similarity).
FT   CONFLICT     44     44       R -> W (in Ref. 1).
SQ   SEQUENCE   345 AA;  39324 MW;  C1D1D8B6BDEA00A0 CRC64;

  Query Match             84.3%;  Score 1525.5;  DB 1;  Length 345;
  Best Local Similarity   84.1%;  Pred. No. 8.8e-103;
  Matches  292;  Conservative   25;  Mismatches   21;  Indels    9;  Gaps    2;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              |||||:|||||| :|: |||:|| |||||||| | ||| |::||| | |||||||: ||
Db          1 MAQKENAYPWPYGSKTSQSGLNTLSQRVLRKEPATTSALALVNRSNSQSTAAPGQKLAEN 60

Qy         61 -------SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQI 113
                  | |: :   : ||||:||||||||||||||||||||| ||||||:||||||
Db         61 KSQGSTASQGSQN--KQPFTIDNFEIGRPLGKGKFGNVYLAREKKSRFIVALKILFKSQI 118

Qy        114 EKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFD 173
              |||||||||||||||||| ||||||:|||||||::|||||||||||||||||||| |||
Db        119 EKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPRGELYKELQKSRTFD 178

Qy        174 EQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTM 233
              ||||||||||:||| ||| |||||||||||||||:||||||||||||||||||||||||
Db        179 EQRTATIMEELSDALTYCHKKKVIHRDIKPENLLLGLQGELKIADFGWSVHAPSLRRKTM 238

Qy        234 CGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFP 293
              |||||||||||||||||| ||||||||||||:|||||||| ||:|||||||||||||||
Db        239 CGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHSETYRRIVKVDLKFP 298

Qy        294 ASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              :|||:||||||||||:||| :|||||:|:||||||||||||||||||
Db        299 SSVPSGAQDLISKLLKHNPWQRLPLAEVAAHPWVRANSRRVLPPSAL 345


RESULT 5
AURB_RAT
ID   AURB_RAT       STANDARD;      PRT;   343 AA.
AC   O55099;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase 12 (EC 2.7.1.37) (Aurora- and Ipl1-
DE   like midbody-associated protein 1) (AIM-1) (Aurora-B).
GN   Name=Aurkb; Synonyms=Aim1, Stk12;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=98119810; PubMed=9450992; DOI=10.1093/emboj/17.3.667;
RA   Terada Y., Tatsuka M., Suzuki F., Yasuda Y., Fujita S., Otsu M.;
RT   "AIM-1: a mammalian midbody-associated protein required for
RT   cytokinesis.";
RL   EMBO J. 17:667-676(1998).
CC   -!- FUNCTION: May be directly involved in regulating the cleavage of
CC       polar spindle microtubules and is a key regulator for the onset of
CC       cytokinesis during mitosis (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localized to the midzone of central spindle
CC       in late anaphase and concentrated into the midbody in telophase
CC       and cytokinesis. Colocalized with gamma tubulin in the mid-body
CC       (By similarity).
CC   -!- TISSUE SPECIFICITY: High level expression seen in the testis. It
CC       is also expressed in the spleen, lung and heart. Expressed in the
CC       G2/M phase of the cell cycle.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D89731; BAA23794.1; -.
DR   HSSP; P31751; 1GZK.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       80    330       Protein kinase.
FT   NP_BIND      86     94       ATP (By similarity).
FT   BINDING     109    109       ATP (By similarity).
FT   ACT_SITE    203    203       Proton acceptor (By similarity).
SQ   SEQUENCE   343 AA;  39234 MW;  27B740D20E287598 CRC64;
```

```
  Query Match             84.1%;  Score 1520.5;  DB 1;  Length 343;
  Best Local Similarity   84.5%;  Pred. No. 2e-102;
  Matches  290;  Conservative   20;  Mismatches   30;  Indels    3;  Gaps    1;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              |||||| |||||| :|: |||:||||||||||||    |  ||:||| | || |||| || 
Db          1 MAQKENVYPWPYGSKTSQSGLNTLPQRVLRKEPAVTPAQALMNRSNSQSTAVPGQKLTEN 60

Qy         61 SSGTP---DILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEG 117
                 |         : ||||:||||||||||||||||||||| ||||||:|||||||||| 
Db         61 KGATALQGSQSRQPFTIDNFEIGRPLGKGKFGNVYLAREKKSRFIVALKILFKSQIEKEG 120

Qy        118 VEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRT 177
              ||||||||||||||| |||||:|||||||::|||||||||||||||||||||| ||||||| 
Db        121 VEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPRGELYKELQKSGTFDEQRT 180

Qy        178 ATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
              |||||||:||||||| |||||||||||||||:|||||||||||||||||||||||||||| 
Db        181 ATIMEELSDALMYCHKKKVIHRDIKPENLLLGLQGELKIADFGWSVHAPSLRRKTMCGTL 240

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              |||||||||||||||| |||||||||:|||||||| ||:|||||||||||||:|:|| 
Db        241 DYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHSETYRRIVKVDLKFPSSMP 300

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              ||:|||||||:||||:|||| ||||||||||||||||||||| 
Db        301 LGAKDLISKLLKHNPSQRLPLEQVSAHPWVRANSRRVLPPSAL 343


RESULT 6
Q8C6C1
ID   Q8C6C1      PRELIMINARY;      PRT;   345 AA.
AC   Q8C6C1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus ES cells cDNA, RIKEN full-length enriched library,
DE   clone:2410042H09 product:serine/threonine kinase 5, full insert
DE   sequence.
GN   Name=Aurkb;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
RA   Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AK075951; BAC36078.1; -.
DR   HSSP; O14965; 1OL6.
DR   MGD; MGI:107168; Aurkb.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   345 AA;  39384 MW;  D204D8B91CFF00A4 CRC64;

  Query Match             84.0%;  Score 1519.5;  DB 2;  Length 345;
  Best Local Similarity   83.9%;  Pred. No. 2.4e-102;
  Matches  291;  Conservative   25;  Mismatches   22;  Indels    9;  Gaps    2;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||:|||||| :|: |||:|| |||||||| | ||| |::| | | |||||||: || 
Db          1 MAQKENAYPWPYGSKTSQSGLNTLSQRVLRKEPATTSALALVNRFNSQSTAAPGQKLAEN 60

Qy         61 -------SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQI 113
                     | |: :   : ||||:||||||||||||||||||||| ||||||:|||||| 
Db         61 KSQGSTASQGSQN--KQPFTIDNFEIGRPLGKGKFGNVYLAREKKSRFIVALKILFKSQI 118

Qy        114 EKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFD 173
              |||||||||||||||| |||||:|||||||::|||||||||||||||||||||| ||| 
Db        119 EKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPRGELYKELQKSRTFD 178

Qy        174 EQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTM 233
              |||||||||||:||| ||| |||||||||||||||||:|||||||||||||||||||||| 
Db        179 EQRTATIMEELSDALTYCHKKKVIHRDIKPENLLLGLQGELKIADFGWSVHAPSLRRKTM 238
```

```
Qy        234 CGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFP 293
              ||||||||||||||||||| ||||||||||||:|||||||| ||:||||||||||||||
Db        239 CGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHSETYRRIVKVDLKFP 298

Qy        294 ASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              :|||:||||||||||:||| :|||||:|:|||||||||||||||||
Db        299 SSVPSGAQDLISKLLKHNPWQRLPLAEVAAHPWVRANSRRVLPPSAL 345


RESULT 7
Q6DE08
ID   Q6DE08      PRELIMINARY;      PRT;   361 AA.
AC   Q6DE08;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   LOC398457 protein.
GN   Name=LOC398457;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RA   Klein S., Strausberg R.;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC077339; AAH77339.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   361 AA;  41735 MW;  384803072040015B CRC64;

  Query Match             68.5%;  Score 1239;  DB 2;  Length 361;
  Best Local Similarity   68.9%;  Pred. No. 6.1e-82;
  Matches  244;  Conservative   40;  Mismatches   52;  Indels   18;  Gaps    6;

Qy          1 MAQKENSYPWPY-GRQTAPSGLSTLPQRVLRKEP-----VTPSALVLMSR---SNVQPTA 51
              |: |||  |  |  : | ||  :|  ||||||||      |||  :|  |   | : |:|
Db          1 MSYKENLNPSSYTSKFTTPSS-ATAAQRVLRKEPYVSTFTTPSDNLLAQRTQLSRITPSA 59

Qy         52 A---PGQ-----KVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIV 103
              :   ||:     ::   ::   ::  | ||||||:||||||||||||||||||:: ||:
Db         60 SSSVPGRVAVSTEMPSQNTALAEMPKRKFTIDDFDIGRPLGKGKFGNVYLAREKQNKFIM 119

Qy        104 ALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELY 163
              ||||||||:|||||||||||||||:|| ||||||:||||:||:||||:||:|||||||
Db        120 ALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNYFHDRKRIYLMLEFAPRGELY 179

Qy        164 KELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSV 223
              |||||   |||||:|| |||||||| ||| :|||||||||||:| ||||||||||||
Db        180 KELQKHGRFDEQRSATFMEELADALHYCHERKVIHRDIKPENLLMGYKGELKIADFGWSV 239

Qy        224 HAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYR 283
              |||||||:|||||||||||||||: |:||||||| |||||| ||| |||:| || ||:|
Db        240 HAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLCYEFLVGMPPFDSPSHTETHR 299

Qy        284 RIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
              ||| ||||||  :  |::|||||||::| :||||  |  ||||:|||||||||
Db        300 RIVNVDLKFPPFLSDGSKDLISKLLRYHPPQRLPLKGVMEHPWVKANSRRVLPP 353


RESULT 8
Q7ZYT9
ID   Q7ZYT9      PRELIMINARY;      PRT;   371 AA.
AC   Q7ZYT9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LOC398457 protein (Fragment).
GN   Name=LOC398457;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
```

```
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC041288; AAH41288.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT   NON_TER       1       1
SQ   SEQUENCE   371 AA;  42716 MW;  19F0C009D0E38CFC CRC64;

  Query Match             68.5%;  Score 1239;  DB 2;  Length 371;
  Best Local Similarity   68.9%;  Pred. No. 6.4e-82;
  Matches  244;  Conservative   40;  Mismatches   52;  Indels   18;  Gaps    6;

Qy          1 MAQKENSYPWPY-GRQTAPSGLSTLPQRVLRKEP-----VTPSALVLMSR---SNVQPTA 51
              |: |||  |  |  : | ||  :|  ||||||||       |||  :|  |    | : |:|
Db         11 MSYKENLNPSSYTSKFTTPSS-ATAAQRVLRKEPYVSTFTTPSDNLLAQRTQLSRITPSA 69

Qy         52 A---PGQ-----KVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIV 103
              :   ||:     ::   ::   ::  | ||||||:||||||||||||||||||:: ||:
Db         70 SSSVPGRVAVSTEMPSQNTALAEMPKRKFTIDDFDIGRPLGKGKFGNVYLAREKQNKFIM 129

Qy        104 ALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELY 163
              |||||||||:||||||||||||||:|| ||||||:||||:||:||||:||:|||||||
Db        130 ALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNYFHDRKRIYLMLEFAPRGELY 189

Qy        164 KELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSV 223
              |||||   |||||:|| |||||||| ||| :|||||||||||||:| ||||||||||||
Db        190 KELQKHGRFDEQRSATFMEELADALHYCHERKVIHRDIKPENLLMGYKGELKIADFGWSV 249

Qy        224 HAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYR 283
              |||||||:|||||||||||||||: |:|||||| |||||| ||| ||||:| || ||:|
Db        250 HAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLCYEFLVGMPPFDSPSHTETHR 309

Qy        284 RIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
              ||| ||||||  :  |::|||||||::| :||||  |  ||||:|||||||||
Db        310 RIVNVDLKFPPPFLSDGSKDLISKLLRYHPPQRLPLKGVMEHPWVKANSRRVLPP 363


RESULT 9
Q9DF70
ID   Q9DF70      PRELIMINARY;      PRT;   361 AA.
AC   Q9DF70;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Protein kinase AIRK2.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20453769; PubMed=10996078; DOI=10.1016/S0960-9822(00)00673-4;
RA   Adams R.R., Wheatleya S.P., Gouldsworthy A.M., Kandels-Lewis S.E.,
RA   Carmena M., Smythe C., Gerloff D.L., Earnshaw W.C.;
RT   "INCENP binds the Aurora-related kinase AIRK2 and is required to
RT   target it to chromosomes, the central spindle and cleavage furrow.";
RL   Curr. Biol. 10:1075-1078(2000).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AF292096; AAG10787.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   361 AA;  41777 MW;  BE577A9135FD0C93 CRC64;

  Query Match             68.4%;  Score 1238;  DB 2;  Length 361;
  Best Local Similarity   68.9%;  Pred. No. 7.3e-82;
  Matches  244;  Conservative   39;  Mismatches   53;  Indels   18;  Gaps    6;

Qy          1 MAQKENSYPWPY-GRQTAPSGLSTLPQRVLRKEP-----VTPSALVLMSR---SNVQPTA 51
              |  |||  |  |  : | ||  :|  ||||||||       |||  :|  |    | : |:|
Db          1 MEYKENLNPSSYTSKFTTPSS-ATAAQRVLRKEPYVSTFTTPSDNLLAQRTQLSRITPSA 59

Qy         52 A---PGQ-----KVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIV 103
              :   ||:     ::   ::   ::  | ||||||:||||||||||||||||||:: ||:
Db         60 SSSVPGRVAVSTEMPSQNTALAEMPKRQFTIDDFDIGRPLGKGKFGNVYLAREKQNKFIM 119

Qy        104 ALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELY 163
              |||||||||:||||||||||||||:|| ||||||:||||:||:||||:||:|||||||
Db        120 ALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNYFHDRKRIYLMLEFAPRGELY 179

Qy        164 KELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSV 223
              |||||   |||||:|| |||||||| ||| :|||||||||||||:| ||||||||||||
Db        180 KELQKHGRFDEQRSATFMEELADALHYCHERKVIHRDIKPENLLMGYKGELKIADFGWSV 239

Qy        224 HAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYR 283
              |||||||:|||||||||||||||: |:|||||| |||||| ||| ||||:| || ||:|
Db        240 HAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLCYEFLVGMPPFDSPSHTETHR 299

Qy        284 RIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
              ||| ||||||  :  |::|||||||::| :||||  |  ||||:|||||||||
Db        300 RIVNVDLKFPPPFLSDGSKDLISKLLRYHPPQRLPLKGVMEHPWVKANSRRVLPP 353


RESULT 10
Q8JG74
ID   Q8JG74      PRELIMINARY;      PRT;   361 AA.
AC   Q8JG74;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Aurora B.
```

```
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22209728; PubMed=12221116;
RA   Bolton M.A., Lan W., Powers S.E., McCleland M.L., Kuang J.,
RA   Stukenberg P.T.;
RT   "Aurora B kinase exists in a complex with survivin and INCENP and its
RT   kinase activity is stimulated by survivin binding and
RT   phosphorylation.";
RL   Mol. Biol. Cell 13:3064-3077(2002).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY115554; AAM76715.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   361 AA;  41777 MW;  1FB77EF41E1D07E8 CRC64;

  Query Match             68.4%;  Score 1237;  DB 2;  Length 361;
  Best Local Similarity   68.9%;  Pred. No. 8.6e-82;
  Matches  244;  Conservative   39;  Mismatches   53;  Indels   18;  Gaps    6;

Qy          1 MAQKENSYPWPY-GRQTAPSGLSTLPQRVLRKEP-----VTPSALVLMSR---SNVQPTA 51
              |  |||  |  |  : | ||  :|  ||||||||      |||  :|  |    | : |:|
Db          1 MEYKENLNPSSYTSKFTTPSS-ATAAQRVLRKEPYVSTFTTPSDNLLAQRTQLSRITPSA 59

Qy         52 A---PGQ-----KVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIV 103
              :   ||:     ::   ::   ::  | ||||||:||||||||||||||||||:: ||:
Db         60 SSSVPGRVAVSTEMPSQNTALAEMPKRKFTIDDFDIGRPLGKGKFGNVYLAREKQNKFIM 119

Qy        104 ALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELY 163
              |||||||||:|||||||||||||||:|| ||||||:||||:||:||||:|:|||||||
Db        120 ALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNYFHDRKRIYLMLEFAPRGELY 179

Qy        164 KELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSV 223
              |||||   |||||:|| |||||||| ||| :||||||||||||:| ||||||||||||
Db        180 KELQKHGRFDEQRSATFMEELADALHYCHERKVIHRDIKPENLLMGYKGELKIADFGWSV 239

Qy        224 HAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYR 283
              |||||||:|||||||||||||||: |:|||||| |||||| ||| |||:| || ||:|
Db        240 HAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLCYEFLVGMPPFDSPSHTETHR 299

Qy        284 RIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
              ||| ||||||  :  |::||||||||::| :||||  |  ||||:|||||||||
Db        300 RIVNVDLKFPPPFLSDGSKDLISKLLRYHPPQRLPLKGVMEHPWVKANSRRVLPP 353


RESULT 11
Q6GPL3
ID   Q6GPL3      PRELIMINARY;      PRT;   368 AA.
AC   Q6GPL3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC83575 protein.
GN   Name=MGC83575;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocytes;
RA   Klein S., Gerhard D.S.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC073103; AAH73103.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   368 AA;  42015 MW;  D1E79FB995701214 CRC64;

  Query Match             68.2%;  Score 1234.5;  DB 2;  Length 368;
  Best Local Similarity   66.4%;  Pred. No. 1.3e-81;
  Matches  239;  Conservative   40;  Mismatches   58;  Indels   23;  Gaps    3;

Qy          1 MAQKENSYPWPYGRQTAPS------GLSTLPQRVLRKEP-----VTPSALVLMSRSNVQP 49
              |: |||  |      ::  |       :|  ||||||:|      |||  :|  |: : |
Db          1 MSYKENLIPSSCSSSSSSSSKFATPSSATAAQRVLRKQPYVSIFTTPSDNLLAQRAQLAP 60

Qy         50 TAAP-----------GQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREK 97
```

```
                |           |  |  :::   :   | ||||||:||||||||||||||||||||
Db         61 RVTPAASSSVPGRVALGTDVASHNTALAEAPKRKFTIDDFDIGRPLGKGKFGNVYLAREK 120

Qy         98 KSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYA 157
              :: ||:|||||||||:|||||||||||||||:|| ||||||:||||:||:||||:||:|
Db        121 QNKFIMALKVLFKSQLEKEGVEHQLRREIEIQSHLRHPNILRMYNYFHDRKRIYLMLEFA 180

Qy        158 PRGELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIA 217
              |||||||||||   |||||:|| |||||||| ||| :|||||||||||||:| |||||||
Db        181 PRGELYKELQKHGRFDEQRSATFMEELADALQYCHERKVIHRDIKPENLLMGYKGELKIA 240

Qy        218 DFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESAS 277
              |||||||||||||:|||||||||||||||||: |:||||||| |||||| ||| |||:| |
Db        241 DFGWSVHAPSLRRRTMCGTLDYLPPEMIEGKTHDEKVDLWCAGVLCYEFLVGMPPFDSPS 300

Qy        278 HNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
              |:||:|||| ||||||  :  |::|||||||::|::||||  |  ||||:|||||||||
Db        301 HSETHRRIVNVDLKFPPPFLSEGSKDLISKLLRYHPAQRLPLKGVMEHPWVKANSRRVLPP 360


RESULT 12
AURC_HUMAN
ID   AURC_HUMAN     STANDARD;      PRT;   309 AA.
AC   Q9UQB9; O60681; O75442; Q9UPK5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Serine/threonine-protein kinase 13 (EC 2.7.1.37) (Aurora/Ipl1/Eg2
DE   protein 2) (Aurora/Ipl1-related kinase 3) (Aurora-C).
GN   Name=AURKC; Synonyms=AIE2, AIK3, STK13;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Testis;
RX   MEDLINE=99025616; PubMed=9809744;
RA   Tseng T.-C., Chen S.-H., Hsu Y.-P.P., Tang T.K.;
RT   "Protein kinase profile of sperm and eggs: cloning and
RT   characterization of two novel testis-specific protein kinases (AIE1,
RT   AIE2) related to yeast and fly chromosome segregation regulators.";
RL   DNA Cell Biol. 17:823-833(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Placenta;
RX   MEDLINE=99017981; PubMed=9799611; DOI=10.1006/geno.1998.5522;
RA   Bernard M., Sanseau P., Henry C., Couturier A., Prigent C.;
RT   "Cloning of STK13, a third human protein kinase related to Drosophila
RT   aurora and budding yeast Ipl1 that maps on chromosome 19q13.3-ter.";
RL   Genomics 53:406-409(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Testis;
RX   MEDLINE=99167501; PubMed=10066797; DOI=10.1074/jbc.274.11.7334;
RA   Kimura M., Matsuda Y., Yoshioka T., Okano Y.;
RT   "Cell cycle-dependent expression and centrosome localization of a
RT   third human Aurora/Ipl1-related protein kinase, AIK3.";
RL   J. Biol. Chem. 274:7334-7340(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   PubMed=15057824; DOI=10.1038/nature02399;
RA   Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J.,
RA   Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M.,
RA   Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E.,
RA   Caenepeel S., Carrano A.V., Caoile C., Chan Y.M., Christensen M.,
RA   Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Detter J.C.,
RA   Escobar J., Flowers D., Fotopulos D., Garcia C., Georgescu A.M.,
RA   Glavina T., Gomez M., Gonzales E., Groza M., Hammon N., Hawkins T.,
RA   Haydu L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H.,
RA   Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,
RA   Malfatti S., Martinez D., McCready P.M., Medina C., Morgan J.,
RA   Nelson K., Nolan M., Ovcharenko I., Pitluck S., Pollard M.,
RA   Popkie A.P., Predki P., Quan G., Ramirez L., Rash S., Retterer J.,
RA   Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D.,
RA   Slezak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A.,
RA   Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I.,
RA   Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E.,
RA   Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,
RA   Rubin E.M., Lucas S.M.;
RT   "The DNA sequence and biology of human chromosome 19.";
RL   Nature 428:529-535(2004).
CC   -!- FUNCTION: May play a part in organizing microtubules in relation
CC       to the function of the centrosome/spindle pole during mitosis.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localized to centrosome from anaphase to
CC       cytokinesis.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9UQB9-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9UQB9-2; Sequence=VSP_004872;
CC   -!- TISSUE SPECIFICITY: Expression is limited to testis. Elevated
CC       expression levels were seen only in a subset of cancer cells such
CC       as HepG2, HuH7 and HeLa cells. Expression is maximum at M phase.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF054621; AAC25955.1; -.
DR   EMBL; AF059681; AAC77369.1; -.
DR   EMBL; AB017332; BAA76292.1; -.
DR   EMBL; AC005261; -; NOT_ANNOTATED_CDS.
DR   HSSP; P49137; 1KWP.
DR   Genew; HGNC:11391; AURKC.
DR   MIM; 603495; -.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Alternative splicing; ATP-binding; Cell cycle;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       43    293       Protein kinase.
FT   NP_BIND      49     57       ATP (By similarity).
FT   BINDING      72     72       ATP (By similarity).
FT   ACT_SITE    166    166       Proton acceptor (By similarity).
FT   VARSPLIC      1     34       Missing (in isoform 2).
FT                                /FTId=VSP_004872.
FT   CONFLICT    109    109       Y -> H (in Ref. 1).
FT   CONFLICT    150    150       L -> V (in Ref. 2).
FT   CONFLICT    193    195       SLR -> LPE (in Ref. 2).
SQ   SEQUENCE   309 AA;  35591 MW;  41B7DFCA91704201 CRC64;

  Query Match             68.0%;  Score 1230;  DB 1;  Length 309;
  Best Local Similarity   76.7%;  Pred. No. 2.3e-81;
  Matches  237;  Conservative   27;  Mismatches   35;  Indels   10;  Gaps    3;

Qy         35 TPSALVLMSRSNVQPT----AAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGN 90
              :| |:| : ::  ||     |   |   : ||       |  |:||||||||||||||||
Db          3 SPRAVVQLGKA--QPAGEELATANQTAQQPSSPA----MRRLTVDDFEIGRPLGKGKFGN 56

Qy         91 VYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRI 150
```

```
                |||||  |:|||||||||||||||||||:|||||||||||||| ||||||||||:| ||:
Db         57 VYLARLKESHFIVALKVLFKSQIEKEGLEHQLRREIEIQAHLQHPNILRLYNYFHDARRV 116

Qy        151 YLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGL 210
                |||||||||||||||||    ||||||||:||||||| ||| ||||||||||||||||
Db        117 YLILEYAPRGELYKELQKSEKLDEQRTATIIEELADALTYCHDKKVIHRDIKPENLLLGF 176

Qy        211 KGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGN 270
                :||:|||||||||| |||||||||||||||||||||| ::||||||||||||||||||
Db        177 RGEVKIADFGWSVHTPSLRRKTMCGTLDYLPPEMIEGRTYDEKVDLWCIGVLCYELLVGY 236

Qy        271 PPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRAN 330
                ||||||||:||||||:|||::|| |:| ||:||||:|||: | |||||||:  ||||:|:
Db        237 PPFESASHSETYRRILKVDVRFPLSMPLGARDLISRLLRYQPLERLPLAQILKHPWVQAH 296

Qy        331 SRRVLPPSA 339
                ||||||| |
Db        297 SRRVLPPCA 305


RESULT 13
Q6AZY8
ID   Q6AZY8      PRELIMINARY;      PRT;   289 AA.
AC   Q6AZY8;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Director MGC Project;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC075064; AAH75064.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
FT   NON_TER       1      1
SQ   SEQUENCE   289 AA;  33542 MW;  70303465BB285F7D CRC64;

  Query Match             67.6%;  Score 1222;  DB 2;  Length 289;
  Best Local Similarity   84.4%;  Pred. No. 8.1e-81;
  Matches  228;  Conservative   21;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
                |  |:|||||||||||||||||||| |:||||||||||||||||||:||||||||||
Db         16 RRLTVDDFEIGRPLGKGKFGNVYLARLKESHFIVALKVLFKSQIEKEGLEHQLRREIEIQ 75

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
                ||| ||||||||||:| ||:||||||||||||||||    ||||||||:|||||||
Db         76 AHLQHPNILRLYNYFHDARRVYLILEYAPRGELYKELQKSEKLDEQRTATIIEELADALT 135

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
                ||| |||||||||||||||| :||:|||||||||| ||||||||||||||||||||||||
Db        136 YCHDKKVIHRDIKPENLLLGFRGEVKIADFGWSVHTPSLRRKTMCGTLDYLPPEMIEGRT 195

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLR 309
                ::|||||||||||||||||| ||||||||:||||||:|||::|| |:| ||:||||:|||
Db        196 YDEKVDLWCIGVLCYELLVGYPPFESASHSETYRRILKVDVRFPLSMPLGARDLISRLLR 255

Qy        310 HNPSERLPLAQVSAHPWVRANSRRVLPPSA 339
                : | |||||||:  ||||:|:|||||||| |
Db        256 YQPLERLPLAQILKHPWVQAHSRRVLPPCA 285


RESULT 14
Q6DLZ0
ID   Q6DLZ0      PRELIMINARY;      PRT;   290 AA.
AC   Q6DLZ0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Aurora/Ipl1-related kinase 3 transcript variant 1.
GN   Name=AURKC;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yan X., Cao L., Yu L.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY661554; AAT64422.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   290 AA;  33672 MW;  D07FD6988B208140 CRC64;
```

```
  Query Match             67.6%;  Score 1222;  DB 2;  Length 290;
  Best Local Similarity   84.4%;  Pred. No. 8.1e-81;
  Matches  228;  Conservative   21;  Mismatches   21;  Indels    0;  Gaps    0;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              |  |:||||||||||||||||||||| |:||||||||||||||||||:|||||||||||
Db         17 RRLTVDDFEIGRPLGKGKFGNVYLARLKESHFIVALKVLFKSQIEKEGLEHQLRREIEIQ 76

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              ||| |||||||||||:| ||:||||||||||||||||||    ||||||||:|||||||
Db         77 AHLQHPNILRLYNYFHDARRVYLILEYAPRGELYKELQKSEKLDEQRTATIIEELADALT 136

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||| |||||||||||||||| :||:|||||||||| ||||||||||||||||||||||
Db        137 YCHDKKVIHRDIKPENLLLGFRGEVKIADFGWSVHTPSLRRKTMCGTLDYLPPEMIEGRT 196

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLR 309
              ::|||||||||||||||||| ||||||||:||||||:|||::|| |:| ||:||||:|||
Db        197 YDEKVDLWCIGVLCYELLVGYPPFESASHSETYRRILKVDVRFPLSMPLGARDLISRLLR 256

Qy        310 HNPSERLPLAQVSAHPWVRANSRRVLPPSA 339
              : | |||||||:  ||||:|:||||||| |
Db        257 YQPLERLPLAQILKHPWVQAHSRRVLPPCA 286


RESULT 15
Q6NW76
ID   Q6NW76      PRELIMINARY;      PRT;   320 AA.
AC   Q6NW76;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Serine/threonine kinase a.
GN   Name=stka;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC067695; AAH67695.1; -.
DR   HSSP; Q63450; 1A06.
DR   ZFIN; ZDB-GENE-020419-40; stka.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   320 AA;  36946 MW;  D1D9EC08556D1CF1 CRC64;

  Query Match             63.9%;  Score 1156;  DB 2;  Length 320;
  Best Local Similarity   66.9%;  Pred. No. 5.7e-76;
  Matches  228;  Conservative   33;  Mismatches   48;  Indels   32;  Gaps    4;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQP----TAAPGQK 56
              |  |||  |                    || :    ||||  |   |  : |     : ||:
Db          1 MQNKENREP----------------RVQQ----TPSAGVGPLRVEMNPDTHAVSGPGRV 39

Qy         57 VMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKE 116
                ::::|       :  :||||:||||||||||||||:|   ::|||||||: ||
Db         40 PVKSNS-------KVLSIDDFDIGRPLGKGKFGNVYLARERKLKVVIALKVLFKSQMVKE 92

Qy        117 GVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQR 176
              |||||||||||||:|| |||||  ||||:|  |::||||||||||:|||||:   ||:||
Db         93 GVEHQLRREIEIQSHLRHPNILRFYNYFHDDTRVFLILEYAPRGEMYKELQRYGRFDDQR 152

Qy        177 TATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGT 236
              ||| |||::||| ||| |||||||||||||||| :|||||||||||||||||:|||||
Db        153 TATYMEEVSDALQYCHEKKVIHRDIKPENLLLGYRGELKIADFGWSVHAPSLRRRTMCGT 212

Qy        237 LDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASV 296
              ||||||||||  |:|||||| |||||| |||||||:||| |||:|| ||||:||  |
Db        213 LDYLPPEMIEGHSHDEKVDLWSIGVLCYECLVGNPPFETASHAETYKRITKVDLQFPKLV 272

Qy        297 PTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
                ||:|||||||||:|| ||||  |  ||||:|||||||||
Db        273 SEGARDLISKLLRHSPSMRLPLRSVMEHPWVKANSRRVLPP 313


RESULT 16
AURC_MOUSE
ID   AURC_MOUSE     STANDARD;      PRT;   282 AA.
AC   O88445; Q9JLC2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase 13 (EC 2.7.1.37) (Aurora/Ipl1/Eg2
DE   protein 1) (Aurora-C).
GN   Name=Aurkc; Synonyms=AIE1, Stk13;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=99025616; PubMed=9809744;
RA   Tseng T.C., Chen S.H., Hsu Y.P., Tang T.K.;
RT   "Protein kinase profile of sperm and eggs: cloning and
RT   characterization of two novel testis-specific protein kinases (AIE1,
RT   AIE2) related to yeast and fly chromosome segregation regulators.";
```

```
RL   DNA Cell Biol. 17:823-833(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129;
RX   MEDLINE=20551164; PubMed=11098217; DOI=10.1089/10445490050199063;
RA   Hu H.M., Chuang C.K., Lee M.J., Tseng T.C., Tang T.K.;
RT   "Genomic organization, expression, and chromosome localization of a
RT   third aurora-related kinase gene, Aie1.";
RL   DNA Cell Biol. 19:679-688(2000).
CC   -!- FUNCTION: May play a part in organizing microtubules in relation
CC       to the function of the centrosome/spindle pole during mitosis.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localized to spindle pole from anaphase to
CC       cytokinesis (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed only in testis.
CC   -!- DEVELOPMENTAL STAGE: Expressed on day 14 dpc in prepubertal
CC       testis, expression reached its plateau on day 21 dpc and remained
CC       at a high level in adult.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF054620; AAC25954.1; -.
DR   EMBL; AF195272; AAF25838.1; -.
DR   HSSP; P31751; 1GZK.
DR   MGD; MGI:1321119; Aurkc.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Serine/threonine-protein kinase; Transferase.
FT   DOMAIN       16    266       Protein kinase.
FT   NP_BIND      22     30       ATP (By similarity).
FT   BINDING      45     45       ATP (By similarity).
FT   ACT_SITE    139    139       Proton acceptor (By similarity).
FT   CONFLICT     74     74       R -> P (in Ref. 2).
SQ   SEQUENCE   282 AA;  32907 MW;  7CB8AEC4984B7883 CRC64;

  Query Match            63.8%;  Score 1154;  DB 1;  Length 282;
  Best Local Similarity  78.2%;  Pred. No. 6.8e-76;
  Matches  219;  Conservative   26;  Mismatches   31;  Indels    4;  Gaps    1;

Qy         58 MENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEG 117
              || |: |     :||||:|||||||||:|||| ||||| |::|||||||||||:|||||
Db          1 MEPSTST----RKHFTINDFEIGRPLGRGKFGRVYLARLKENHFIVALKVLFKSEIEKEG 56

Qy        118 VEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRT 177
              :|||||||:|||||| | |||||||||||  |||||||||| ||||||||:    |:|||
Db         57 LEHQLRREVEIQAHLQHRNILRLYNYFYDDTRIYLILEYAPGGELYKELQRHQKLDQQRT 116

Qy        178 ATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
              |||::||:||| ||| |||||||||||||||| ||:||:||||||| ||||||||||||
Db        117 ATIIQELSDALTYCHEKKVIHRDIKPENLLLGLNGEVKISDFGWSVHTPSLRRKTMCGTL 176

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              ||||||||  : :|| |||||||||||||||| |||||:: :|||||| :|| |||:|||
Db        177 DYLPPEMIAQKPYNEMVDLWCIGVLCYELLVGKPPFESSTSSETYRRIRQVDFKFPSSVP 236

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
              ||||||||||::||||| ||||  ||||| :|||||||
Db        237 AGAQDLISKLLRYHPSERLSLAQVLKHPWVREHSRRVLPP 276


RESULT 17
Q6P209
ID   Q6P209      PRELIMINARY;      PRT;   276 AA.
AC   Q6P209;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Aurkc protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Unfertilized egg;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Unfertilized egg;
RA   Strausberg R.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC064780; AAH64780.1; -.
DR   HSSP; Q63450; 1A06.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   276 AA;  32358 MW;  91530FF97594D63E CRC64;

  Query Match            63.4%;  Score 1147;  DB 2;  Length 276;
  Best Local Similarity  78.1%;  Pred. No. 2.1e-75;
  Matches  218;  Conservative   26;  Mismatches   31;  Indels    4;  Gaps    1;

Qy         58 MENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEG 117
              || |: |     :||||:|||||||||:|||| ||||| |::|||||||||||:|||||
```

```
Db          1 MEPSTST----RKHFTINDFEIGRPLGRGKFGRVYLARLKENHFIVALKVLFKSEIEKEG 56

Qy        118 VEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRT 177
              :|||||||:|||||| | ||||||||||  |||||||||| ||||||||:    |:|||
Db         57 LEHQLRREVEIQAHLQHRNILRLYNYFYDDTRIYLILEYAPGGELYKELQRHQKLDQQRT 116

Qy        178 ATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
              |||::||:||| ||| |||||||||||||||||| ||:||:||||||| ||||||||||
Db        117 ATIIQELSDALTYCHEKKVIHRDIKPENLLLGLNGEVKISDFGWSVHTPSLRRKTMCGTL 176

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              ||||||||  : :|| |||||||||||||||| |||||:: :|||||| :|| |||:|||
Db        177 DYLPPEMIAQKPYNEMVDLWCIGVLCYELLVGKPPFESSTSSETYRRIRQVDFKFPSSVP 236

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLP 336
              ||||||||||::||||| ||||  ||||| :||||||
Db        237 AGAQDLISKLLRYHPSERLSLAQVLKHPWVREHSRRVLP 275


RESULT 18
Q8JGS8
ID   Q8JGS8      PRELIMINARY;      PRT;   346 AA.
AC   Q8JGS8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Aurora-like serine/threonine kinase (Fragment).
GN   Name=stka;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22035902; PubMed=12006978; DOI=10.1038/ng896;
RA   Golling G., Amsterdam A., Sun Z., Antonelli M., Maldonado E., Chen W.,
RA   Burgess S., Haldi M., Artzt K., Farrington S., Lin S.-Y., Nissen R.M.,
RA   Hopkins N.;
RT   "Insertional mutagenesis in zebrafish rapidly identifies genes
RT   essential for early vertebrate development.";
RL   Nat. Genet. 31:135-140(2002).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY099518; AAM28206.1; -.
DR   HSSP; O14965; 1OL6.
DR   ZFIN; ZDB-GENE-020419-40; stka.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT   NON_TER       1       1
SQ   SEQUENCE   346 AA;  39977 MW;  C8A2CEC221C7C2D9 CRC64;

  Query Match           61.4%;  Score 1111.5;  DB 2;  Length 346;
  Best Local Similarity   73.2%;  Pred. No. 1.1e-72;
  Matches  210;  Conservative   31;  Mismatches   39;  Indels    7;  Gaps    1;

Qy         51 AAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFK 110
              : ||:  ::::|       :  :||||:||||||||||||||||:|   ::|||||||
Db         60 SGPGRVPVKSNS-------KVLSIDDFDIGRPLGKGKFGNVYLARERKLKVVIALKVLFK 112

Qy        111 SQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSC 170
              ||: ||||||||||||||:|| |||||| ||||:|  |::||||||||||:|||||:
Db        113 SQMVKEGVEHQLRREIEIQSHLRHPNILRFYNYFHDDTRVFLILEYAPRGEMYKELQRYG 172

Qy        171 TFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRR 230
              ||:||||| |||::||| ||| ||||||||||||||| :|||||||||||||||||||||
Db        173 HFDDQRTATYMEEVSDALQYCHEKKVIHRDIKPENLLLGYRGELKIADFGWSVHAPSLRR 232

Qy        231 KTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDL 290
              :|||||||||||||||  |:|||||| ||||||| ||||||||:  | |||:|| ||||
Db        233 RTMCGTLDYLPPEMIEGHSHDEKVDLWSIGVLCYECLVGNPPFETRQHAETYKRITKVDL 292

Qy        291 KFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
              :||  |  ||:|||||||||:|| ||||  |  |  |:|||||||||
Db        293 QFPKLVSEGARDLISKLLRHSPSMRLPLRSVMEHRGVKANSRRVLPP 339


RESULT 19
STK6_MOUSE
ID   STK6_MOUSE      STANDARD;      PRT;   395 AA.
AC   P97477; O35624; Q91YU4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase 6 (EC 2.7.1.37) (Aurora-family kinase
DE   1) (Aurora/IPL1-related kinase 1) (Ipl1- and aurora-related kinase 1)
DE   (Aurora-A) (Serine/threonine kinase Ayk1).
GN   Name=Stk6; Synonyms=Ark1, Ayk1, Iak1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=BALB/c; TISSUE=Testis;
RX   MEDLINE=97392693; PubMed=9245792; DOI=10.1083/jcb.138.3.643;
RA   Gopalan G., Chan C.S.M., Donovan P.J.;
RT   "A novel mammalian, mitotic spindle-associated kinase is related to
RT   yeast and fly chromosome segregation regulators.";
RL   J. Cell Biol. 138:643-656(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Testis;
RX   MEDLINE=97349097; PubMed=9205101; DOI=10.1038/sj.onc.1201144;
RA   Yanai A., Arama E., Kilfin G., Motro B.;
RT   "ayk1, a novel mammalian gene related to Drosophila aurora centrosome
RT   separation kinase, is specifically expressed during meiosis.";
RL   Oncogene 14:2943-2950(1997).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   STRAIN=BALB/c;
RX   MEDLINE=98183439; PubMed=9514916; DOI=10.1006/bbrc.1998.8250;
RA   Shindo M., Nakano H., Kuroyanagi H., Shirasawa T., Mihara M.,
RA   Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;
RT   "cDNA cloning, expression, subcellular localization, and chromosomal
RT   assignment of mammalian aurora homologues, aurora-related kinase (ARK)
RT   1 and 2.";
RL   Biochem. Biophys. Res. Commun. 244:285-292(1998).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
```

```
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   REVIEW.
RX   MEDLINE=21306577; PubMed=11413462; DOI=10.1038/35048096;
RA   Nigg E.A.;
RT   "Mitotic kinases as regulators of cell division and its checkpoints.";
RL   Nat. Rev. Mol. Cell Biol. 2:21-32(2001).
CC   -!- FUNCTION: May play a role in cell cycle regulation during anaphase
CC       and/or telophase, in relation to the function of the
CC       centrosome/spindle pole region during chromosome segregation.
CC       Maybe involved in microtubule formation and/or stabilization.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes on centrosomes in interphase cells
CC       and at each spindle pole in mitosis (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P97477-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P97477-2; Sequence=VSP_004871;
CC         Note=May be less abundant or less stable;
CC   -!- TISSUE SPECIFICITY: Highly expressed in testis, in about one third
CC       of the seminiferous tubules. Expression is restricted to specific
CC       spermatocytes nearing completion of prophase, with levels falling
CC       off on transition to elongated spermatids. Highly expressed in the
CC       ovary, expression in the oocyte starts around the transition to
CC       large growing follicle. Abundant expression is seen in the
CC       proliferating granulosa and thecal cells of the growing follicle,
CC       and in the young corpus luteum. Very weakly expressed in spleen
CC       and intestine.
CC   -!- DEVELOPMENTAL STAGE: At 7.5-9.5 dpc expressed evenly all over the
CC       embryo. At later stages, expression is mainly restricted to
CC       proliferating zones. The highest levels of expression at mid-
CC       embryonic development (13.5 dpc) were observed in the liver, lung,
CC       kidney and back (trapezius) muscle and all regions in active
CC       proliferation.
CC   -!- PTM: Phosphorylated (By similarity).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U80932; AAB62982.1; -.
DR   EMBL; AF007817; AAB63205.1; -.
DR   EMBL; U69106; AAC12682.1; -.
DR   EMBL; BC005425; AAH05425.1; -.
DR   EMBL; BC014711; AAH14711.1; -.
DR   HSSP; P31751; 1GZK.
DR   MGD; MGI:894678; Stk6.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Alternative splicing; ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      124    374       Protein kinase.
FT   NP_BIND     130    138       ATP (By similarity).
FT   BINDING     153    153       ATP (By similarity).
FT   ACT_SITE    247    247       Proton acceptor (By similarity).
FT   VARSPLIC      1      1       M -> MAVEGEPGCCKRIGKAVWRRGDM (in isoform
FT                                2).
FT                                /FTId=VSP_004871.
FT   CONFLICT    234    234       A -> T (in Ref. 1).
SQ   SEQUENCE   395 AA;  44772 MW;  26B6B65105A1A812 CRC64;

  Query Match             59.9%;  Score 1084;  DB 1;  Length 395;
  Best Local Similarity   62.7%;  Pred. No. 1.3e-70;
  Matches  210;  Conservative   43;  Mismatches   78;  Indels    4;  Gaps    2;

Qy          6 NSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSS--G 63
              ||   |   |   :|    |:::      | : :   : | || |:       | :|
Db         51 NSQRVPSQAQKLGAGQKPAPKQLPAASVPRPVSRLNNPQKNEQPAASGNDSEKEQASLQK 110

Qy         64 TPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
              | |   | :|::||:|||||||||||||||::| ||:|||||||:|:||  ||||||
Db        111 TEDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLR 170

Qy        124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
              ||:|||:|| |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |
Db        171 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITE 230

Qy        184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
              ||:|| ||| |:|||||||||||||  ||||||||||||||| || ||||||||||||
Db        231 LANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPE 290

Qy        244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
              ||||||:|||||| :|||||| ||| |||| :: |||||| :|:  ||  |  ||:||
Db        291 MIEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDL 350

Qy        304 ISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPS 338
              ||:||:|| |:|| ||:|  |||::||| :  ||:
Db        351 ISRLLKHNASQRLTLAEVLEHPWIKANSSK--PPT 383


RESULT 20
Q7TNK2
ID   Q7TNK2      PRELIMINARY;      PRT;   395 AA.
AC   Q7TNK2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Serine/threonine protein kinase 6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RA   Zhou G., Li W., Yu L.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY336976; AAQ16152.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
```

```
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   395 AA;  44771 MW;  271DC321221A8DD9 CRC64;

  Query Match           59.9%;  Score 1084;  DB 2;  Length 395;
  Best Local Similarity   62.7%;  Pred. No. 1.3e-70;
  Matches  210;  Conservative   43;  Mismatches   78;  Indels   4;  Gaps    2;

Qy          6 NSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSS--G 63
              ||   |   |   :|    |:::       | : :   : | || |:      | :|
Db         51 NSQRVPSQAQKLGAGQKPAPKQLPAASVPRPVSRLNNPQKNEQPAASGNDSEKEQASLQK 110

Qy         64 TPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
              | |   | :|::||:||||||||||||||||||::| ||:|||||||:|:||  ||||||
Db        111 TEDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLR 170

Qy        124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
              ||:|||:|| |||||||| ||:|  |:||||||| | :|:|||   |||||||| : |
Db        171 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITE 230

Qy        184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
              ||:|| ||| |:|||||||||||||  |||||||||||||| || ||||||||||||
Db        231 LANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPE 290

Qy        244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
              |||||||:|||||| :|||||| ||| |||| : :: |||||| :|:  ||  |  ||:||
Db        291 MIEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDL 350

Qy        304 ISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPS 338
              ||:||:|| |:|| ||:|  |||::||| :  ||:
Db        351 ISRLLKHNASQRLTLAEVLEHPWIKANSSK--PPT 383


RESULT 21
Q8C3H8
ID   Q8C3H8      PRELIMINARY;      PRT;   386 AA.
AC   Q8C3H8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched
DE   library, clone:D830019E15 product:serine/threonine kinase 6, full
DE   insert sequence.
GN   Name=Stk6;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Heart;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AK085861; BAC39557.1; -.
DR   HSSP; O14965; 1OL6.
DR   MGD; MGI:894678; Stk6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   386 AA;  43968 MW;  31970E9CDE31C6A7 CRC64;

  Query Match           59.7%;  Score 1080;  DB 2;  Length 386;
  Best Local Similarity   63.0%;  Pred. No. 2.4e-70;
  Matches  208;  Conservative   42;  Mismatches   78;  Indels    2;  Gaps    1;

Qy          6 NSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSS--G 63
              ||   |   |   :|    |:::       | : :   : | || |:      | :|
Db         51 NSQRVPSQAQKLGAGQKPAPKQLPAASVPRPVSRLNNPQKNEQPAASGNDSEKEQASLQK 110

Qy         64 TPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
              | |   | :|::||:||||||||||||||||||::| ||:|||||||:|:||  ||||||
Db        111 TEDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLR 170

Qy        124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
              ||:|||:|| |||||||| ||:|  |:||||||| | :|:|||   |||||||| : |
Db        171 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITE 230
```

```
Qy        184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
              ||:|| ||| |:|||||||||||||||  |||||||||||||||| || ||||||||||||
Db        231 LANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPE 290

Qy        244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
              |||||||:|||||| :|||||| ||| |||||: :: |||||| :|:  ||  |  ||:||
Db        291 MIEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDL 350

Qy        304 ISKLLRHNPSERLPLAQVSAHPWVRANSRR 333
              ||:||:|| |:|| ||:|  |||::||| :
Db        351 ISRLLKHNASQRLTLAEVLEHPWIKANSSK 380


RESULT 22
Q8BP87
ID   Q8BP87      PRELIMINARY;      PRT;   395 AA.
AC   Q8BP87;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched
DE   library, clone:5730433I06 product:serine/threonine kinase 6, full
DE   insert sequence.
GN   Name=Stk6;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AK077513; BAC36838.1; -.
DR   HSSP; O14965; 1OL6.
DR   MGD; MGI:894678; Stk6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   395 AA;  44813 MW;  3B081DFE3DC52667 CRC64;

  Query Match             59.3%;  Score 1073;  DB 2;  Length 395;
  Best Local Similarity   62.1%;  Pred. No. 7.9e-70;
  Matches  208;  Conservative   43;  Mismatches   80;  Indels    4;  Gaps    2;

Qy          6 NSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSS--G 63
              ||   |   |  :|    |:::       | : :   : | || |:       | :|
Db         51 NSQRVPSQAQKLGAGQKPAPKQLPAASVPRPVSRLNNPQKNEQPAASGNDSEKEQASLQK 110

Qy         64 TPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
              | |   | :|::||:|||||||||||||||::| ||:|||||||:|:||  ||||||
Db        111 TEDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLR 170

Qy        124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
              ||:|||:|| |||||||| ||:|  |:|||||||| | :|:||||   |||||| | : |
Db        171 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTPTYITE 230

Qy        184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
              ||:|| ||| |:|||||||||||||  |||||||||||||||| || ||||||||||||
Db        231 LANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPE 290

Qy        244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
              |||||||:|||||| :|||||  || |||| :: |||||| :|:  ||  |  ||:||
Db        291 MIEGRMHDEKVDLWSLGVLCYEFQVGMPPFEAHTYQETYRRISRVEFTFPDFVTEGARDL 350

Qy        304 ISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPS 338
              ||:||:|| |:|| ||:|  |||::||| :  ||:
Db        351 ISRLLKHNASQRLTLAEVLEHPWIKANSSK--PPT 383


RESULT 23
STK6_RAT
ID   STK6_RAT       STANDARD;      PRT;   397 AA.
AC   P59241;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
```

```
DE   Serine/threonine-protein kinase 6 (EC 2.7.1.37) (Aurora-A) (ratAurA).
GN   Name=Stk6; Synonyms=Stk15;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar Furth; TISSUE=Mammary gland;
RX   MEDLINE=22119444; PubMed=12124350;
RA   Goepfert T.M., Adigun Y.E., Zhong L., Gay J., Medina D.,
RA   Brinkley W.R.;
RT   "Centrosome amplification and overexpression of aurora A are early
RT   events in rat mammary carcinogenesis.";
RL   Cancer Res. 62:4115-4122(2002).
RN   [2]
RP   REVIEW.
RX   MEDLINE=21306577; PubMed=11413462; DOI=10.1038/35048096;
RA   Nigg E.A.;
RT   "Mitotic kinases as regulators of cell division and its checkpoints.";
RL   Nat. Rev. Mol. Cell Biol. 2:21-32(2001).
CC   -!- FUNCTION: May play a role in cell cycle regulation during anaphase
CC       and/or telophase, in relation to the function of the
CC       centrosome/spindle pole region during chromosome segregation.
CC       Maybe involved in microtubule formation and/or stabilization. May
CC       play a key role during tumor development and progression.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes on centrosomes in interphase cells
CC       and at each spindle pole in mitosis (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in mammary gland and tumor.
CC   -!- INDUCTION: Activated by progesterone.
CC   -!- PTM: Phosphorylated (By similarity).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF537333; AAN06823.1; -.
DR   HSSP; P31751; 1GZK.
DR   RGD; 628895; Stk6.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      126    376       Protein kinase.
FT   NP_BIND     132    140       ATP (By similarity).
FT   BINDING     155    155       ATP (By similarity).
FT   ACT_SITE    249    249       Proton acceptor (By similarity).
SQ   SEQUENCE   397 AA;  44874 MW;  95DECA2198DCED85 CRC64;

  Query Match           59.2%;  Score 1071;  DB 1;  Length 397;
  Best Local Similarity  60.8%;  Pred. No. 1.1e-69;
  Matches  211;  Conservative   39;  Mismatches   77;  Indels   20;  Gaps    3;

Qy         14 RQTAPSGLSTLPQR----------VLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSS- 62
              |   ||    :| :          ||:: |          || | : |       |:|
Db         45 RVLCPSNSQRVPPQAQKPVAGQKPVLKQLPAASGPRPASRLSNPQKSEQPQPAASGNNSE 104

Qy         63 -------GTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEK 115
                  | |   | :|::||:||||||||||||||||||:| ||:||||||| |:||
Db        105 KEQTSIQKTEDSKKRQWTLEDFDIGRPLGKGKFGNVYLAREKQSKFILALKVLFKVQLEK 164

Qy        116 EGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQ 175
              |||||||||:|||:|| |||||||| ||:|  |:|||||||| | :|:||||    ||||
Db        165 AGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQ 224

Qy        176 RTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCG 235
              |||| : |||:|| ||| |:|||||||||||||  |||||||||||||||| || |:||
Db        225 RTATYITELANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTLCG 284

Qy        236 TLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS 295
              |||| |||||||||||:|||||| :|||||| ||| ||||: :: |||||| :|:  ||
Db        285 TLDYQPPEMIEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETYRRISRVEFTFPDF 344

Qy        296 VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              |  ||:||||:||:|| |:|| ||:|  |||::||| :  ||:   |
Db        345 VTEGARDLISRLLKHNSSQRLTLAEVLEHPWIKANSSK--PPTGHNS 389


RESULT 24
ST6L_XENLA
ID   ST6L_XENLA     STANDARD;      PRT;   408 AA.
AC   Q91819;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase Eg2-like (EC 2.7.1.37) (p46X1Eg22).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Egg;
RA   Roghi C., Le Guellec R., Paris J., Couturier A., Philippe M.;
RT   "Eg2, selected by differential screening encodes a new Xenopus protein
RT   kinase family.";
RL   Submitted (OCT-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis
CC       especially from prophase through anaphase. Partially colocalised
CC       with gamma tubulin in the centrosome, from S to M phase (By
CC       similarity).
CC   -!- PTM: Phosphorylated (By similarity).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z17206; CAA78914.1; ALT_INIT.
DR   HSSP; P06244; 1FOT.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      140    390       Protein kinase.
FT   NP_BIND     146    154       ATP (By similarity).
```

```
FT   BINDING     169    169       ATP (By similarity).
FT   ACT_SITE    263    263       Proton acceptor (By similarity).
SQ   SEQUENCE   408 AA;  46461 MW;  97F6A69C7357AEE8 CRC64;

  Query Match             58.7%;  Score 1062;  DB 1;  Length 408;
  Best Local Similarity   59.8%;  Pred. No. 5.2e-69;
  Matches  207;  Conservative   44;  Mismatches   63;  Indels   32;  Gaps    4;

Qy         20 GLSTLPQRVL--RKEPV----TPSALVLM----------------SRSNVQPTAAPGQKV 57
              | | :||||:   ::||     |:|  |:                    | | |:
Db         51 GPSNVPQRVMMQAQKPVLSNQKPTAQGLLRPATHGHQTSKPQGPNENRNPQQTSHSSTPN 110

Qy         58 MENSSGTPDILT----------RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKV 107
              ||     |    |          : : ::|||||||||||||||||||::| ||:||||
Db        111 MEKKGSTDQGKTLAVPKEEGKKKQWCLEDFEIGRPLGKGKFGNVYLARERESKFILALKV 170

Qy        108 LFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQ 167
              |||||:|| |||||||||:|||:|| |||||||| ||:|  |:||||:||| |||::|||
Db        171 LFKSQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDASRVYLILDYAPGGELFRELQ 230

Qy        168 KSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPS 227
              |   ||:||:|  :::||:||:||| ||||||||||||||  ||||||||||||||||
Db        231 KCTRFDDQRSALYIKQLAEALLYCHSKKVIHRDIKPENLLLGSNGELKIADFGWSVHAPS 290

Qy        228 LRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVK 287
               || |:||||||||||||||:|||||| :|||||| ||| |||| :| |||||| |
Db        291 SRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFETDTHQETYRRISK 350

Qy        288 VDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRR 333
              |: ::|  |   |:||:||||:|||: |||| |  |||:  ||::
Db        351 VEFQYPPYVSEEAKDLVSKLLKHNPNHRLPLKGVLEHPWIVKNSQQ 396


RESULT 25
Q6DJK0
ID   Q6DJK0      PRELIMINARY;      PRT;   408 AA.
AC   Q6DJK0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   LOC398349 protein.
GN   Name=LOC398349;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Klein S., Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC075177; AAH75177.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   408 AA;  46477 MW;  FD4855CC6708E5E7 CRC64;

  Query Match             58.6%;  Score 1060;  DB 2;  Length 408;
  Best Local Similarity   59.8%;  Pred. No. 7.2e-69;
  Matches  207;  Conservative   44;  Mismatches   63;  Indels   32;  Gaps    4;

Qy         20 GLSTLPQRVL--RKEPV----TPSALVLM----------------SRSNVQPTAAPGQKV 57
              | | :||||:   ::||     |:|  |:                    | | |:
Db         51 GPSNVPQRVMMQAQKPVLSNQKPTAQGLLRPATHGHQTSKPQGPNENRNPQQTSHSSTPN 110

Qy         58 MENSSGTPDILT----------RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKV 107
              ||     |    |          : : ::|||||||||||||||||||::| ||:||||
Db        111 MEKKGSTDQGKTLAVPKEEGKKKQWCLEDFEIGRPLGKGKFGNVYLARERESKFILALKV 170

Qy        108 LFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQ 167
              |||||:|| |||||||||:|||:|| |||||||| ||:|  |:||||:||| |||::|||
Db        171 LFKSQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDACRVYLILDYAPGGELFRELQ 230

Qy        168 KSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPS 227
              |   ||:||:|  :::||:||:||| ||||||||||||||  ||||||||||||||||
Db        231 KCTRFDDQRSALYIKQLAEALLYCHSKKVIHRDIKPENLLLGSNGELKIADFGWSVHAPS 290

Qy        228 LRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVK 287
               || |:||||||||||||||:|||||| :|||||| ||| |||| :| |||||| |
Db        291 SRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFETDTHQETYRRISK 350

Qy        288 VDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRR 333
              |: ::|  |   |:||:||||:|||: |||| |  |||:  ||::
Db        351 VEFQYPPYVSEEAKDLVSKLLKHNPNHRLPLKGVLEHPWIVKNSQQ 396


RESULT 26
STK6_HUMAN
ID   STK6_HUMAN     STANDARD;      PRT;   403 AA.
AC   O14965; O60445; O75873; Q9BQD6; Q9UPG5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Serine/threonine-protein kinase 6 (EC 2.7.1.37) (Serine/threonine
DE   kinase 15) (Aurora/IPL1-related kinase 1) (Aurora-related kinase 1)
DE   (hARK1) (Aurora-A) (Breast-tumor-amplified kinase).
GN   Name=STK6; Synonyms=AIK, ARK1, AURA, BTAK, STK15;
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=97298083; PubMed=9153231; DOI=10.1074/jbc.272.21.13766;
RA   Kimura M., Kotani S., Hattori T., Sumi N., Yoshioka T., Todokoro K.,
RA   Okano Y.;
RT   "Cell cycle-dependent expression and spindle pole localization of a
RT   novel human protein kinase, Aik, related to Aurora of Drosophila and
RT   yeast Ipl1.";
RL   J. Biol. Chem. 272:13766-13771(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98183439; PubMed=9514916; DOI=10.1006/bbrc.1998.8250;
RA   Shindo M., Nakano H., Kuroyanagi H., Shirasawa T., Mihara M.,
RA   Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;
RT   "cDNA cloning, expression, subcellular localization, and chromosomal
RT   assignment of mammalian aurora homologues, aurora-related kinase (ARK)
RT   1 and 2.";
RL   Biochem. Biophys. Res. Commun. 244:285-292(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast;
RX   MEDLINE=98442657; PubMed=9771714; DOI=10.1038/2496;
RA   Zhou H., Kuang J., Zhong L., Kuo W.-L., Gray J.W., Sahin A.,
RA   Brinkley B.R., Sen S.;
RT   "Tumour amplified kinase STK15/BTAK induces centrosome amplification,
RT   aneuploidy and transformation.";
RL   Nat. Genet. 20:189-193(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Wang L., Thibodeau S.N.;
RT   "Mutational analysis of the STK15 gene in human tumors.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052; DOI=10.1038/414865a;
RA   Deloukas P., Matthews L.H., Ashurst J.L., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cervix, Colon, Kidney, and Muscle;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   CELL-CYCLE REGULATION.
RX   MEDLINE=21895866; PubMed=11790771; DOI=10.1074/jbc.M108252200;
RA   Tanaka M., Ueda A., Kanamori H., Ideguchi H., Yang J., Kitajima S.,
RA   Ishigatsubo Y.;
RT   "Cell-cycle-dependent regulation of human aurora A transcription is
RT   mediated by periodic repression of E4TF1.";
RL   J. Biol. Chem. 277:10719-10726(2002).
RN   [8]
RP   REVIEW.
RX   MEDLINE=21306577; PubMed=11413462; DOI=10.1038/35048096;
RA   Nigg E.A.;
RT   "Mitotic kinases as regulators of cell division and its checkpoints.";
RL   Nat. Rev. Mol. Cell Biol. 2:21-32(2001).
CC   -!- FUNCTION: May play a role in cell cycle regulation during anaphase
CC       and/or telophase, in relation to the function of the
CC       centrosome/spindle pole region during chromosome segregation.
CC       Maybe involved in microtubule formation and/or stabilization. May
CC       play a key role during tumor development and progression.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes on centrosomes in interphase cells
CC       and at each spindle pole in mitosis.
CC   -!- TISSUE SPECIFICITY: Highly expressed in testis and weakly in
CC       skeletal muscle, thymus and spleen. Also highly expressed in
CC       colon, ovarian, prostate, neuroblastoma, breast and cervical
CC       cancer cell lines. Expression is cell-cycle regulated, low in
CC       G1/S, accumulates during G2/M, and decreases rapidly after.
CC   -!- PTM: Phosphorylated.
CC   -!- DISEASE: Defects in STK6 are responsible for numerical centrosome
CC       aberrations including aneuploidy.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in positions 105, 125, 129, 235 and 241.
CC   -!- CAUTION: Although authors have considered STK6 and STK15 as two
CC       different proteins, it is clear that they are the same protein.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D84212; BAA23592.1; ALT_FRAME.
DR   EMBL; AF008551; AAC12708.1; -.
DR   EMBL; AF011467; AAC23448.1; -.
DR   EMBL; AF011468; AAC63902.1; -.
DR   EMBL; AF195947; AAF29508.1; -.
DR   EMBL; AF195942; AAF29508.1; JOINED.
DR   EMBL; AF195943; AAF29508.1; JOINED.
DR   EMBL; AF195944; AAF29508.1; JOINED.
DR   EMBL; AF195945; AAF29508.1; JOINED.
DR   EMBL; AF195946; AAF29508.1; JOINED.
```

```
DR   EMBL; AL121914; CAC12717.1; -.
DR   EMBL; BC001280; AAH01280.1; -.
DR   EMBL; BC002499; AAH02499.1; -.
DR   EMBL; BC006423; AAH06423.1; -.
DR   EMBL; BC027464; AAH27464.1; -.
DR   PDB; 1MQ4; X-ray; A=120-391.
DR   PDB; 1MUO; X-ray; A=107-403.
DR   PDB; 1OL6; X-ray; A=122-403.
DR   PDB; 1OL7; X-ray; A=122-403.
DR   Genew; HGNC:11393; STK15.
DR   Genew; HGNC:11409; STK6.
DR   H-InvDB; HIX0015930; -.
DR   MIM; 602687; -.
DR   MIM; 603072; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005819; C:spindle; TAS.
DR   GO; GO:0004672; F:protein kinase activity; TAS.
DR   GO; GO:0007067; P:mitosis; TAS.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   3D-structure; ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      133    383       Protein kinase.
FT   NP_BIND     139    147       ATP (By similarity).
FT   BINDING     162    162       ATP (By similarity).
FT   ACT_SITE    256    256       Proton acceptor (By similarity).
FT   CONFLICT     31     31       F -> I (in Ref. 3).
FT   CONFLICT     57     57       V -> I (in Ref. 2).
SQ   SEQUENCE   403 AA;  45809 MW;  125F3594834CD157 CRC64;

  Query Match             58.5%;  Score 1057.5;  DB 1;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.1e-68;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 27
STK6_XENLA
ID   STK6_XENLA     STANDARD;      PRT;   407 AA.
AC   Q91820;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Serine/threonine-protein kinase Eg2 (EC 2.7.1.37) (pEg2) (p46Eg265).
GN   Name=EG2;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Egg;
RX   MEDLINE=98122855; PubMed=9454730;
RA   Roghi C., Giet R., Uzbekov R., Morin N., Chartrain I., Le Guellec R.,
RA   Couturier A., Doree M., Phillippe M., Prigent C.;
RT   "The Xenopus protein kinase pEg2 associates with the centrosome in a
RT   cell cycle-dependent manner, binds to the spindle microtubules and is
RT   involved in bipolar mitotic spindle assembly.";
RL   J. Cell Sci. 111:557-572(1998).
CC   -!- FUNCTION: Associates with the centrosome in a cell-cycle dependent
CC       manner and invades the microtubules at the poles of the spindle
CC       during mitosis suggesting that it may be involved in the corect
CC       formation of bipolar mitotic spindles.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis
CC       especially from prophase through anaphase. Partially colocalised
CC       with gamma tubulin in the centrosome, from S to M phase.
CC   -!- TISSUE SPECIFICITY: Highly expressed in ovary and testis.
CC   -!- PTM: Phosphorylated.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Aurora
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z17207; CAA78915.1; -.
DR   PIR; S52243; S52243.
DR   HSSP; Q16539; 1KV2.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Phosphorylation;
KW   Serine/threonine-protein kinase; Transferase.
FT   DOMAIN      140    390       Protein kinase.
FT   NP_BIND     146    154       ATP (By similarity).
FT   BINDING     169    169       ATP (By similarity).
FT   ACT_SITE    263    263       Proton acceptor (By similarity).
SQ   SEQUENCE   407 AA;  46372 MW;  DE1628A2C6D11277 CRC64;

  Query Match             58.1%;  Score 1050.5;  DB 1;  Length 407;
  Best Local Similarity   56.3%;  Pred. No. 3.5e-68;
  Matches  205;  Conservative   46;  Mismatches   71;  Indels   42;  Gaps    4;

Qy         11 PYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSR---SNVQPT-------AAPGQKVME- 59
              |   |  | | :   ||:|    |   |   :   |: :||       |  | :  :
Db         32 PQSTQFRPPGTAVSAQRILGPSNVPQRVLAQAQKPILSSQKPTTQIPLRPATQGHQSSKP 91

Qy         60 ------------NSSGTPDI------------------LTRHFTIDDFEIGRPLGKGKF 88
                          : | ||::                : : ::||||||||||||||
Db         92 QGPNENRNPQQTSHSSTPNVEKKGSTDQGKTSAVPKEEGKKKQWCLEDFEIGRPLGKGKF 151
```

```
Qy          89 GNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRR 148
               ||||||||::| ||:|||||||||:|| |||||||||:|||:|| |||||||| ||:|
Db         152 GNVYLARERESKFILALKVLFKSQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDAS 211

Qy         149 RIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLL 208
               |:||||:||| |||::|||   ||:||:|  :::||:||:||| ||||||||||||||
Db         212 RVYLILDYAPGGELFRELQKCTRFDDQRSAMYIKQLAEALLYCHSKKVIHRDIKPENLLL 271

Qy         209 GLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLV 268
               |  |||||||||||||| || |:|||||||||||||||:| |||| :|||||| ||
Db         272 GSNGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDETVDLWSLGVLCYEFLV 331

Qy         269 GNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVR 328
               | ||||: :| |||||| ||: ::|  |   |:||:||||:|||: ||||  |  |||:
Db         332 GKPPFETDTHQETYRRISKVEFQYPPYVSEEARDLVSKLLKHNPNHRLPLKGVLEHPWII 391

Qy         329 ANSR 332
                ||:
Db         392 KNSQ 395


RESULT 28
Q6DBZ4
ID   Q6DBZ4      PRELIMINARY;      PRT;   405 AA.
AC   Q6DBZ4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Zgc:100912 protein.
GN   Name=zgc:100912;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Singapore local strain; TISSUE=Embryo;
RX   PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Singapore local strain; TISSUE=Embryo;
RA   Director MGC Project;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC078304; AAH78304.1; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   405 AA;  45447 MW;  B04793FC827BD489 CRC64;

  Query Match             57.6%;  Score 1041.5;  DB 2;  Length 405;
  Best Local Similarity   58.4%;  Pred. No. 1.6e-67;
  Matches  195;  Conservative   49;  Mismatches   79;  Indels   11;  Gaps    2;

Qy          11 PYGRQTAPSGLSTLPQR---VLRKEPVTPSALVLMSRSNVQP--------TAAPGQKVME 59
               | |: : |   :| |::       :|      |   :    :|         | :
Db          61 PVGKTSCPGDQNTRPEQHKPAAHSKPQPKPPLSAETNKTAEPSKQDKPHQTPSSTTSTSS 120

Qy          60 NSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVE 119
               |:||:     : :|:::|:||| |||||||:||||||::: ||:|||||||| |:|| |||
Db         121 NTSGSSKSSKKAWTLENFDIGRALGKGKFGSVYLAREQQTKFILALKVLFKKQLEKAGVE 180

Qy         120 HQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTAT 179
               ||||||:|||:|| |||||||| ||:|  |:|||||:||:|||| |||:  |||:||:||
Db         181 HQLRREVEIQSHLRHPNILRLYGYFHDAARVYLILEFAPKGELYGELQRCGTFDDQRSAT 240

Qy         180 IMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDY 239
                : |||||| ||| | |||||||||||||  |||||||||||||| || || |:||||||
Db         241 YIMELADALRYCHSKAVIHRDIKPENLLLGANGELKIADFGWSVHTPSSRRSTLCGTLDY 300

Qy         240 LPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTG 299
               ||||||||: |:|||||| :|||||| ||| |||| || ||||:| :|:  :|| |  |
Db         301 LPPEMIEGKTHDEKVDLWSLGVLCYEFLVGRPPFETKSHEETYRKISRVEFTYPAHVSNG 360

Qy         300 AQDLISKLLRHNPSERLPLAQVSAHPWVRANSRR 333
               ::|||::||:|||  |||:  |  ||||  || :
Db         361 SRDLINRLLKHNPMHRLPIQGVMEHPWVVENSTK 394


RESULT 29
Q9M077
ID   Q9M077      PRELIMINARY;      PRT;   294 AA.
AC   Q9M077;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative serine/threonine protein kinase.
GN   Name=AT4g32830;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lennard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
RA   Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AL161582; CAB80000.1; -.
DR   PIR; T10690; T10690.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
```

```
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   294 AA;  33972 MW;  B6669E1799083BE4 CRC64;

  Query Match           52.0%;  Score 940;  DB 2;  Length 294;
  Best Local Similarity   60.3%;  Pred. No. 2.6e-60;
  Matches  179;  Conservative   47;  Mismatches   63;  Indels    8;  Gaps    4;

Qy         49 PTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVL 108
              ||    |:  | |  :     : :|: ||:||:|||:||||:|||||||:|: :||||||
Db          4 PTETQHQE-KEASDASAAAAQKRWTLSDFDIGKPLGRGKFGHVYLAREKRSNHVVALKVL 62

Qy        109 FKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQK 168
              ||||:::  |||||||:|||:|| |||||||| ||||::|:|||||||| ||||||:|||
Db         63 FKSQLQQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQKRVYLILEYAARGELYKDLQK 122

Qy        169 SCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSL 228
                | |:| || :  || ||:||||| |||||||||||||:| :|||||||||||||  :
Db        123 CKYFSERRAATYVASLARALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN- 181

Qy        229 RRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKV 288
              ||:|||||||||||||:|   |:  ||:| :|:|||| | | ||||:  |::||||||:|
Db        182 RRRTMCGTLDYLPPEMVESVEHDASVDIWSLGILCYEFLYGVPPFEAMEHSDTYRRIVQV 241

Qy        289 DLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
              |||||    :   |:||||::|    |:|||| ::  |||:  |:     || :  |
Db        242 DLKFPPKPIISASAKDLISQMLVKESSQRLPLHKLLEHPWIVQNA----DPSGIYRV 294


RESULT 30
Q8LBX4
ID   Q8LBX4      PRELIMINARY;      PRT;   294 AA.
AC   Q8LBX4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative serine/threonine protein kinase.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22088475; PubMed=12093376;
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AY086942; AAM64506.1; -.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   294 AA;  33926 MW;  E1D048CCF4BEED3F CRC64;

  Query Match           51.8%;  Score 937;  DB 2;  Length 294;
  Best Local Similarity   60.3%;  Pred. No. 4.2e-60;
  Matches  179;  Conservative   46;  Mismatches   64;  Indels    8;  Gaps    4;

Qy         49 PTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVL 108
              ||    |:  | |  :     : :|: ||:||:|||:||||:|||||||:|: :||||||
Db          4 PTETQHQE-KEASDASAAAAQKRWTLSDFDIGKPLGRGKFGHVYLAREKRSNHVVALKVL 62

Qy        109 FKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQK 168
              ||||:::  |||||||:|||:|| |||||||| ||||::|:|||||||| ||||||:|||
Db         63 FKSQLQQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQKRVYLILEYAARGELYKDLQK 122

Qy        169 SCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSL 228
                | |:| || :  || ||:||||| |||||||||||||:| :|||||||||||||  :
Db        123 CKYFSERRAATYVASLARALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFNX 182

Qy        229 RRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKV 288
              || |||||||||||||:|   |:  ||:| :|:|||| | | ||||:  |::||||||:|
Db        183 RR-TMCGTLDYLPPEMVESVEHDASVDIWSLGILCYEFLYGVPPFEAMEHSDTYRRIVQV 241

Qy        289 DLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
              |||||    :   |:||||::|    |:|||| ::  |||:  |:     || :  |
Db        242 DLKFPPKPIISASAKDLISQMLVKESSQRLPLHKLLEHPWIVQNA----DPSGIYRV 294


RESULT 31
O82309
ID   O82309      PRELIMINARY;      PRT;   282 AA.
AC   O82309;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein At2g25880.
GN   Name=At2g25880;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA   Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Town C.D., Kaul S.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Shinn P., Chen H., Cheuk R., Kim C.J., Carninci P., Hayashizaki Y.,
RA   Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,
RA   Seki M., Shinozaki K., Ecker J.R.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AC005395; AAC42257.1; -.
DR   EMBL; BT010653; AAR07517.1; -.
DR   PIR; H84653; H84653.
DR   HSSP; O14965; 1OL6.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
```

```
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   282 AA;  32757 MW;  3475F9C04D166B7E CRC64;

  Query Match            50.9%;  Score 921.5;  DB 2;  Length 282;
  Best Local Similarity   63.7%;  Pred. No. 5.4e-59;
  Matches  174;  Conservative   42;  Mismatches   50;  Indels    7;  Gaps    3;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              : :|  ||:||:|||:||||:|||||||:|  |||||||||:|:::  ||||||||:|||
Db         12 KRWTTSDFDIGKPLGRGKFGHVYLAREKRSDHIVALKVLFKAQLQQSQVEHQLRREVEIQ 71

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| |||||||| ||||::|:|||||| ||||||||||    | |:| || :  || ||:
Db         72 SHLRHPNILRLYGYFYDQKRVYLILEYAVRGELYKELQKCKYFSERRAATYVASLARALI 131

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        132 YCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 190

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|:|||| | | ||||:  |:|||:|||:||||||    | : |:||||::
Db        191 HDASVDIWSLGILCYEFLYGVPPFEAREHSETYKRIVQVDLKFPPKPIVSSSAKDLISQM 250

Qy        308 LRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              |    ::|| | ::  |||:  |:     || |
Db        251 LVKESTQRLALHKLLEHPWIVQNA----DPSGL 279


RESULT 32
Q683C9
ID   Q683C9      PRELIMINARY;      PRT;   288 AA.
AC   Q683C9;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein At2g25880.
GN   Name=At2g25880;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
RA   Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
RA   Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
RA   Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
RA   Hayashizaki Y., Shinozaki K.;
RT   "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK175188; BAD42951.1; -.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   288 AA;  33310 MW;  732F036B2E707909 CRC64;

  Query Match            50.9%;  Score 921.5;  DB 2;  Length 288;
  Best Local Similarity   63.7%;  Pred. No. 5.5e-59;
  Matches  174;  Conservative   42;  Mismatches   50;  Indels    7;  Gaps    3;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              : :|  ||:||:|||:||||:|||||||:|  |||||||||:|:::  ||||||||:|||
Db         18 KRWTTSDFDIGKPLGRGKFGHVYLAREKRSDHIVALKVLFKAQLQQSQVEHQLRREVEIQ 77

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| |||||||| ||||::|:|||||| ||||||||||    | |:| || :  || ||:
Db         78 SHLRHPNILRLYGYFYDQKRVYLILEYAVRGELYKELQKCKYFSERRAATYVASLARALI 137

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        138 YCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 196

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|:|||| | | ||||:  |:|||:|||:||||||    | : |:||||::
Db        197 HDASVDIWSLGILCYEFLYGVPPFEAREHSETYKRIVQVDLKFPPKPIVSSSAKDLISQM 256

Qy        308 LRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              |    ::|| | ::  |||:  |:     || |
Db        257 LVKESTQRLALHKLLEHPWIVQNA----DPSGL 285


RESULT 33
Q9LG71
ID   Q9LG71      PRELIMINARY;      PRT;   432 AA.
AC   Q9LG71;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein P0710E05.26 (Hypothetical protein
DE   P0671B11.2).
GN   Name=P0710E05.26; Synonyms=P0671B11.2;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=12447438; DOI=10.1038/nature01184;
RA   Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA   Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA   Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA   Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA   Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA   Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA   Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA   Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA   Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA   Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K, Shibata M.,
RA   Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA   Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA   Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA   Yano M., Jiang J., Gojobori T.;
RT   "The genome sequence and structure of rice chromosome 1.";
RL   Nature 420:312-316(2002).
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; AP002743; BAA99439.1; -.
DR   EMBL; AP002746; BAB12687.1; -.
DR   HSSP; O14965; 1OL6.
DR   Gramene; Q9LG71; -.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR011009; Kinase_like.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
```

```
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Hypothetical protein; Kinase;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   432 AA;  49787 MW;  42100935A51A5D0F CRC64;

  Query Match             50.1%;  Score 906;  DB 2;  Length 432;
  Best Local Similarity   65.2%;  Pred. No. 1.2e-57;
  Matches  172;  Conservative   40;  Mismatches   48;  Indels    4;  Gaps    3;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              : : : ||:||:|||:||||:||||||||:|: ||||||||||||:::  ||||||||:|||
Db        163 KRWVLSDFDIGKPLGRGKFGHVYLAREKRSNHIVALKVLFKSQLKQSQVEHQLRREVEIQ 222

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| |||||||| |||| |:|||||| :||||||||:   | |:|:|| :  || ||:
Db        223 SHLRHPNILRLYGYFYD-TRVYLILEYALKGELYKELQRCKHFSERRSATYIASLAHALI 281

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              | ||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        282 YLHGKHVIHRDIKPENLLIGSQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVEKTE 340

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|:|||| | | ||||:  |:|||||||||||||    |   |:||||::
Db        341 HDYHVDIWSLGILCYEFLYGVPPFEAKEHSETYRRIVKVDLKFPLKPFVSPAAKDLISQM 400

Qy        308 LRHNPSERLPLAQVSAHPWVRANS 331
              |  | : |||| ::  |||:  |:
Db        401 LVKNSAHRLPLHKLLEHPWIVQNA 424


Search completed: June 20, 2005, 20:12:23
Job time : 124.351 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:46:50 ; Search time 143.505 Seconds
(without alignments)
1086.128 Million cell updates/sec

Title: US-10-734-126-4
Perfect score: 2110
Sequence: 1 MDRSKENCISGPVKATAPVG.....TANSSKPSNCQNKESASKQS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 42

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2110	100.0	403	2	AAW18084	Aaw18084 Human Aur
2	2110	100.0	403	2	AAY22476	Aay22476 Human AUR
3	2110	100.0	420	5	ABP41958	Abp41958 Human ova
4	2109	100.0	403	4	AAG67614	Aag67614 Amino aci
5	2109	100.0	403	4	AAG67435	Aag67435 Amino aci
6	2109	100.0	403	7	ADF61840	Adf61840 Human ser
7	2109	100.0	403	8	ADQ89832	Adq89832 Antagonis
8	2109	100.0	403	8	ADR05174	Adr05174 Human GTP
9	2104	99.7	403	6	ABP97367	Abp97367 Human ser
10	2104	99.7	403	6	ABR48160	Abr48160 Human bla
11	2104	99.7	403	7	ADN39888	Adn39888 Cancer/an
12	2104	99.7	403	8	ADE15337	Ade15337 Truncated
13	2104	99.7	403	8	ADH59556	Adh59556 Monoclona
14	2104	99.7	403	8	ADK67720	Adk67720 Human mod
15	2104	99.7	403	8	ADM72205	Adm72205 Human TAS
16	2103	99.7	403	6	ABP97366	Abp97366 Human ser
17	2095	99.3	403	7	ADJ31727	Adj31727 Human mit
18	1935.5	91.7	402	7	ABR61579	Abr61579 Human HsA
19	1935.5	91.7	402	7	ADD89969	Add89969 Human can
20	1924	91.2	403	5	AAO18740	Aao18740 Human NOV
21	1614	76.5	319	6	ABP97470	Abp97470 Truncated
22	1503	71.2	309	6	ABP97472	Abp97472 Truncated
23	1460	69.2	300	6	ABP97473	Abp97473 Truncated
24	1215	57.6	248	8	ADK71858	Adk71858 Human kin
25	1057.5	50.1	344	2	AAW18083	Aaw18083 Human Aur
26	1057.5	50.1	344	2	AAY22475	Aay22475 Human AUR
27	1057.5	50.1	344	4	AAG67615	Aag67615 Amino aci
28	1057.5	50.1	344	4	AAG67436	Aag67436 Amino aci
29	1057.5	50.1	344	6	ABR48188	Abr48188 Human bla
30	1057.5	50.1	344	6	ABU56641	Abu56641 Lung canc
31	1057.5	50.1	344	6	ABR92152	Abr92152 Human cer
32	1057.5	50.1	344	7	ADB80537	Adb80537 Ovarian c
33	1057.5	50.1	344	7	ADN38885	Adn38885 Cancer/an
34	1057.5	50.1	344	7	ADN95464	Adn95464 Human BEC
35	1057.5	50.1	344	8	ADK67718	Adk67718 Human mod
36	1057.5	50.1	344	8	ADQ89834	Adq89834 Antagonis
37	1057.5	50.1	344	8	ABM81826	Abm81826 Tumour-as
38	1056.5	50.1	344	7	ABR61580	Abr61580 Human HsA
39	1050	49.8	343	2	AAW99783	Aaw99783 Rat AIM-1
40	1046	49.6	347	2	AAY27052	Aay27052 Human pro
41	1046	49.6	347	6	ABU57643	Abu57643 Different
42	1045	49.5	320	8	ADO57331	Ado57331 Kidney de

ALIGNMENTS

RESULT 1
AAW18084
ID AAW18084 standard; protein; 403 AA.
XX
AC AAW18084;
XX
DT 07-SEP-1997 (first entry)
XX
DE Human Aurora-2.
XX
KW Aurora-2; AUR-2; signal transduction; protein kinase; tumour; cancer;
KW protein kinase; gene therapy; diagnosis; antibody.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1. .130
FT /label= N-terminal_domain
FT Misc-difference 50
FT /note= "deduced residue from some cDNA clones is Leu"
FT Misc-difference 50
FT /note= "deduced residue from some cDNA clones is Leu"
FT Misc-difference 57
FT /note= "deduced residue from some cDNA clones is Ile"
FT Misc-difference 57
FT /note= "deduced residue from some cDNA clones is Ile"
FT Domain 131. .403
FT /label= Kinase_domain
FT Modified-site 288
FT /label= Phosphorylation
FT /note= "cAMP-dependent protein kinase phosphorylation
FT site conserved in AUR-2 and yeast and Drosophila
FT homologues"
FT Modified-site 334
FT /label= Phosphorylation
FT /note= "tyrosine phosphorylation consensus site conserved
FT in Drosophila auroura but not in AUR-1 or yeast IPL1"
FT Modified-site 342
FT /label= Phosphorylation
FT /note= "cAMP-dependent protein kinase phosphorylation
FT site conserved in AUR-2 and yeast and Drosophila
FT homologues"
XX
PN WO9722702-A1.
XX
PD 26-JUN-1997.
XX
PF 25-NOV-1996; 96WO-US018859.
XX
PR 18-DEC-1995; 95US-0008809P.
PR 14-AUG-1996; 96US-0023943P.

XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Mossie KG;
XX
DR WPI; 1997-341693/31.
DR N-PSDB; AAT67290.
XX
PT Aurora-1 and Aurora-2 and related genes - useful in tumour gene therapy.
XX
PS Claim 3; Page 83-85; 98pp; English.
XX
CC Novel human proteins, termed Aurora-1 (AAW18083) and Aurora-2 (AAW18084)
CC (AUR-1 and AUR-2), are related serine/threonine kinases with short N-
CC terminal extensions that appear to be involved in cancer and/or signal
CC transduction disorders. Their amino acid sequences were deduced from
CC pancreatic tumour cDNA clones (AAT67289- 90). AUR-1 and AUR-2 appear to
CC regulate nuclear division, with disruption of their signaling resulting
CC in polyploid cells. AUR-2 RNA is low or absent in most normal tissues,
CC and abundant in a subset of tumour-derived cell lines, partic. those of
CC colorectal origin. AUR polypeptides can be expressed in host cells and
CC used to raise diagnostic antibodies and to screen for compounds that
CC interact with AUR-1 and/or AUR-2
XX
SQ Sequence 403 AA;

Query Match 100.0%; Score 2110; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 9.6e-187;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 2
AAY22476
ID AAY22476 standard; protein; 403 AA.
XX
AC AAY22476;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human AUR2 protein sequence.
XX
KW AUR1; AUR2; human; AUR modulator; cancer; glioma; medullablastoma;
KW chondrosarcoma; pancreatic tumour; proliferative disease; diagnosis;
KW therapy.
XX
OS Homo sapiens.
XX
PN WO9937788-A2.
XX
PD 29-JUL-1999.
XX
PF 21-JAN-1999; 99WO-US001283.
XX
PR 22-JAN-1998; 98US-00012135.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Mossie K;
XX
DR WPI; 1999-458699/38.
DR N-PSDB; AAX99725.
XX
PT New nucleic acid encoding human AUR1 and 2 polypeptides, used to identify
PT specific modulators for treating cancer or for diagnosis.
XX
PS Claim 11; Page 141-142; 153pp; English.
XX
CC This sequence is the human AUR2 protein of the invention. The AUR1 and
CC AUR2 proteins can be used to identify specific modulators of, and to
CC generate specific antibodies recognising AUR1 and AUR2. The modulators
CC can be used for treating conditions involving abnormal AUR signal
CC transduction, specifically cancer (of colon, breast, kidney, ovary,
CC bladder, head or neck, also glioma, medullablastoma, chondrosarcoma and
CC pancreatic tumours, particularly of colon (specifically), breast or
CC kidney). The modulators can also be used for studying their effects in
CC animal models of proliferative disease. Probes, based on the coding
CC sequences are used, diagnostically, to detect or quantify AUR mRNA, by
CC hybridisation or polymerase chain reaction (PCR). The DNA, optionally
CC mutated, are useful in gene therapy. Ab are used as diagnostic
CC immunoassay reagents for detecting the proteins
XX
SQ Sequence 403 AA;

Query Match 100.0%; Score 2110; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 9.6e-187;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 3
ABP41958

```
ID   ABP41958 standard; protein; 420 AA.
XX
AC   ABP41958;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HPCOK03, SEQ ID NO:3090.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive; chromosome 20q13.2-13.3.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US018569.
XX
PR   07-JUN-2000; 2000US-0209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   N-PSDB; ABQ55035.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT   cancer), immune disorders, cardiovascular disorders and neurological
PT   diseases.
XX
PS   Claim 11; SEQ ID NO 3090; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
CC   disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC   blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC   respiratory disorders, neurological disorders, gastrointestinal disorders
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds which
CC   modulate ovarian antigen expression or activity. The polynucleotides may
CC   further be used for gene therapy, chromosome mapping, in the
CC   identification of individuals and in forensic analysis, and the
CC   polypeptides may be used as food additives or to prepare antibodies
CC   useful in disease diagnosis, drug targeting and phenotyping. The present
CC   sequence represents a human ovarian antigen of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
```

```
SQ   Sequence 420 AA;

  Query Match           100.0%;  Score 2110;  DB 5;  Length 420;
  Best Local Similarity 100.0%;  Pred. No. 1e-186;
  Matches  403;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         18 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 77

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 137

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 197

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 257

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 317

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 377

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        378 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 420
```

```
RESULT 4
AAG67614
ID   AAG67614 standard; protein; 403 AA.
XX
AC   AAG67614;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human protein.
XX
KW   Human; protein kinase; protein phosphatase; signal transduction.
XX
OS   Homo sapiens.
XX
PN   WO200109316-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-JP005061.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   18-OCT-1999;   99US-0159590P.
PR   11-JAN-2000; 2000JP-00118776.
PR   17-FEB-2000; 2000US-0183322P.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T,  Funahashi S;
PI   Senoo C,  Nezu J;
XX
DR   WPI; 2001-570286/64.
XX
PT   New genes encoding proteins with protein kinase/protein phosphatase
PT   activity, useful in the diagnosis and treatment of diseases.
XX
```

```
PS   Example 4; Page 94-97; 233pp; Japanese.
XX
CC   The specification describes human protein kinase/protein phosphatases. It
CC   is expected that the protein kinase/protein phosphatase gene participates
CC   in signal transduction in cells. The protein kinase/protein phosphatase
CC   polypeptides and polynucleotides are useful for developing diagnostics
CC   and treatment agents for human and animal diseases. The protein
CC   kinase/protein phosphatase polypeptides are useful as target molecules in
CC   designing novel drugs. The protein kinase/protein phosphatase
CC   polynucleotides are useful as a source of probes and primers, which may
CC   be used to isolate homologous sequences. The present sequence represents
CC   a human protein, which is used in the course of the invention
XX
SQ   Sequence 403 AA;

  Query Match             100.0%;  Score 2109;  DB 4;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-186;
  Matches  402;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 5
AAG67435
ID   AAG67435 standard; protein; 403 AA.
XX
AC   AAG67435;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human polypeptide.
XX
KW   Human; protein kinase; protein phosphatase; signal transduction;
KW   intracellular signalling pathway.
XX
OS   Homo sapiens.
XX
PN   WO200109345-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-JP005060.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   18-OCT-1999;   99US-0159590P.
PR   11-JAN-2000; 2000JP-00118776.
PR   17-FEB-2000; 2000US-0183322P.
PR   02-MAY-2000; 2000JP-00183767.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T,  Funahashi S;
PI   Senoo C,  Nezu J;
XX
DR   WPI; 2001-564736/63.
XX
PT   New genes encoding protein kinase and protein phosphatase, useful for
PT   identifying modulators which can be used to treat human or animal
PT   disorders associated with the expression or function of these enzymes.
XX
PS   Example 4; Page 196-199; 336pp; Japanese.
XX
CC   The specification describes human protein kinase/protein phosphatases.
CC   The polypeptides are expected to participate in signal transduction in
CC   cells. The kinase phosphatases are connected with intracellular
CC   signalling pathways. Antisense oligonucleotides and compounds identified
CC   by screening (agonists or antagonists) can be used to treat human or
CC   animal disorders associated with the expression or function of the
CC   protein. In addition, the polypeptides may be used as target molecules
CC   fir drug development. The present sequence represents a polypeptide, used
CC   in the course of the invention
XX
SQ   Sequence 403 AA;

  Query Match             100.0%;  Score 2109;  DB 4;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-186;
  Matches  402;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 6
ADF61840
ID   ADF61840 standard; protein; 403 AA.
XX
AC   ADF61840;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human serine/threonine kinase 15 protein.
XX
KW   cell cycle arrest; cytostatic; antipsoriatic; antiarteriosclerotic;
KW   vasotropic; antithyroid; melanoma; breast; ovarian; lung;
```

```
KW   gastrointestinal; colon cancer; Grave's disease; psoriasis;
KW   atherosclerosis; restenosis; vasoproliferative; human;
KW   serine/threonine kinase 15; STK15; ARK2; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO2003088910-A2.
XX
PD   30-OCT-2003.
XX
PF   15-APR-2003; 2003WO-US011867.
XX
PR   15-APR-2002; 2002US-00123568.
PR   15-APR-2002; 2002US-00123731.
PR   16-APR-2002; 2002US-0373366P.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Hitoshi Y,  Jenkins Y;
XX
DR   WPI; 2003-865396/80.
DR   N-PSDB; ADF61839.
XX
PT   Identifying a compound that modulates cell cycle arrest, for treating
PT   e.g. cancer, comprises contacting a cell comprising a target polypeptide
PT   and determining the chemical or phenotypic effect of the compound upon
PT   the cell.
XX
PS   Claim 1; SEQ ID NO 24; 176pp; English.
XX
CC   The invention relates to a novel method for identifying a compound that
CC   modulates cell cycle arrest comprising contacting a target polypeptide
CC   within a cell with a compound and determining the chemical or phenotypic
CC   effect of the compound upon the cell. The method of the invention has
CC   cytostatic, antipsoriatic, antiarteriosclerotic, vasotropic and
CC   antithyroid applications and may be useful for identifying a compound
CC   that modulates cell cycle arrest. Such compounds may subsequently be used
CC   for developing therapeutic reagents to treat melanoma, breast, ovarian,
CC   lung, gastrointestinal or colon cancer, as well as other proliferative
CC   diseases such as Grave's disease, psoriasis, atherosclerosis, restenosis
CC   and other vasoproliferative diseases. The current sequence is that of the
CC   human serine/threonine kinase 15 (STK15;ARK2) protein of the invention.
XX
SQ   Sequence 403 AA;

  Query Match             100.0%;  Score 2109;  DB 7;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-186;
  Matches  402;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 7
ADQ89832
ID   ADQ89832 standard; protein; 403 AA.
XX
AC   ADQ89832;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Antagonist of cell cycle progression polypeptide #131.
XX
KW   Cytostatic; cancer; cell division cycle; mitosis; meiosis;
KW   cell cycle progression.
XX
OS   Homo sapiens.
XX
PN   WO2004063362-A2.
XX
PD   29-JUL-2004.
XX
PF   31-DEC-2003; 2003WO-GB005635.
XX
PR   10-JAN-2003; 2003US-0439123P.
PR   06-MAY-2003; 2003US-0468402P.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Glover D,  Bell G,  Frenz L,  Midgley C;
XX
DR   WPI; 2004-544089/52.
DR   N-PSDB; ADQ89831.
XX
PT   New cell cycle progression genes and proteins for modulating cell cycle
PT   progression in cells, for preventing, treating or diagnosing cell
PT   proliferative diseases (e.g. cancer) or for identifying modulators of
PT   mitosis or meiosis.
XX
PS   Claim 2; SEQ ID NO 262; 461pp; English.
XX
CC   The present invention relates to a polynucleotide for preventing,
CC   treating or diagnosing a disease in an individual. The composition or the
CC   polypeptide, polynucleotide or RNA precursor, or antibody is useful for
CC   diagnosing, preventing or treating diseases (e.g. cell proliferative
CC   diseases such as cancer) in an individual. These may also be used for
CC   identifying substances capable of binding to or modulating the function
CC   of the polypeptide, capable of affecting the function of the
CC   corresponding gene, or capable of inhibiting the cell division cycle or
CC   cell cycle progression, preferably mitosis and/or meiosis. The present
CC   sequence represents an antagonist of cell cycle progression protein
CC   sequence.
XX
SQ   Sequence 403 AA;

  Query Match             100.0%;  Score 2109;  DB 8;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-186;
  Matches  402;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
```

```
Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 8
ADR05174
ID ADR05174 standard; protein; 403 AA.
XX
AC ADR05174;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human GTPase regulator-associated w focal adhesion kinase pp125 protein.
XX
KW apoptosis; cytostatic; antiinflammatory; antiasthmatic; respiratory;
KW antirheumatic; antiarthritic; gynaecological; cardiant; vasotropic;
KW antipsoriatic; antiulcer; gastrointestinal; immunosuppressive;
KW neuroprotective; cancer; autoimmune; neurodegenerative; inflammatory;
KW asthma; chronic obstructive pulmonary disease; cystic fibrosis;
KW rheumatoid arthritis; acute respiratory distress syndrome; preeclampsia;
KW myocardial ischaemia; reperfusion injury; psoriasis; bronchiolitis;
KW Crohn's disease; ulcerative colitis; inflammatory bowel disease; human;
KW enzyme; GTPase regulator-associated with focal adhesion kinase pp125;
KW GRAF.
XX
OS Homo sapiens.
XX
PN WO2004065959-A2.
XX
PD 05-AUG-2004.
XX
PF 23-JAN-2004; 2004WO-GB000271.
XX
PR 23-JAN-2003; 2003GB-00001566.
PR 25-MAR-2003; 2003US-0457533P.
XX
PA (EIRX-) EIRX THERAPEUTICS LTD.
XX
PI Seery L, Hayes I, Murphy F;
XX
DR WPI; 2004-593556/57.
DR N-PSDB; ADR05175.
XX
PT Identifying a modulator of apoptosis-associated polypeptide function,
PT useful for treating e.g., cancer, comprises incubating a sample
PT containing an apoptosis-associated polypeptide and a candidate agent to
PT permit binding.
XX
PS Claim 1; Page; 230pp; English.
XX
CC The invention relates to a novel method for identifying an agent that
CC modulates the function of an apoptosis-associated polypeptide,
CC particularly a kinase or GCPR (G-protein-coupled receptor). The method
CC comprises providing a sample containing an apoptosis-associated
CC polypeptide and a candidate agent and incubating under conditions to
CC permit binding of the candidate agent to the polypeptide, measuring the
CC binding and comparing it with the binding of the polypeptide to a control
CC agent known not to bind to the polypeptide. The method of the invention
CC has cytostatic, antiinflammatory, antiasthmatic, respiratory,
CC antirheumatic, antiarthritic, gynaecological, cardiant, vasotropic,
CC antipsoriatic, antiulcer, gastrointestinal, immunosuppressive and
CC neuroprotective applications. The method and molecules may be useful for
CC treating a disease or condition characterised by abnormal apoptosis in
CC mammalian tissue, particularly cancer, such as small cell lung cancer,
CC cancer of the kidney, uterus, prostate, bladder, ovary, colon and breast,
CC leukaemias, sarcomas and myelomas. Furthermore, autoimmune,
CC neurodegenerative and inflammatory conditions may be treated, including
CC asthma, chronic obstructive pulmonary disease, cystic fibrosis,
CC rheumatoid arthritis, acute respiratory distress syndrome, preeclampsia,
CC myocardial ischaemia, reperfusion injury, psoriasis, bronchiolitis,
CC Crohn's disease, ulcerative colitis and inflammatory bowel disease. The
CC current sequence is that of a human apoptosis-associated protein of the
CC invention which was used during siRNA (small interfering RNA)-mediated
CC gene silencing.
XX
SQ Sequence 403 AA;

Query Match 100.0%; Score 2109; DB 8; Length 403;
Best Local Similarity 99.8%; Pred. No. 1.2e-186;
Matches 402; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 9
ABP97367
ID ABP97367 standard; protein; 403 AA.
XX
AC ABP97367;
XX
DT 10-MAY-2003 (first entry)
XX
DE Human serine/threonine kinase 15 (STK15), Ile31 variant.
XX
KW Human; serine/threonine kinase 15; STK15; STK6; Aurora2; cell cycle;
KW chromosome 20; centrosome-associated kinase; cancer susceptibility;
KW single nucleotide polymorphism; SNP; genetic diagnosis; prognosis;
KW detection; diagnosis; cancer; malignant astrocytoma; glioblastoma;
KW medulloblastoma; gastric cancer; colorectal cancer; colorectal adenoma;
KW acute myelogenous leukaemia; lung cancer; renal cancer; leukaemia;
KW breast cancer; prostate cancer; endometrial cancer; neuroblastoma;
KW enzyme.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 31

```
FT                   /note= "This residue is replaced by Phe in a polymorphic
FT                   variant"
FT   Misc-difference 57
FT                   /note= "Encoded by ATT"
XX
PN   WO2003012046-A2.
XX
PD   13-FEB-2003.
XX
PF   29-JUL-2002; 2002WO-US024115.
XX
PR   27-JUL-2001; 2001US-0308911P.
PR   28-NOV-2001; 2001US-0334146P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Toland AE,  Balmain A;
XX
DR   WPI; 2003-239517/23.
DR   N-PSDB; ABZ97367.
XX
PT   Determining cancer susceptibility in a human subject comprises
PT   identifying in a nucleic acid sample from the subject, a nucleotide
PT   occurrence of a single polynucleotide polymorphism (SNP) of the STK15
PT   gene.
XX
PS   Claim 63; Page; 92pp; English.
XX
CC   The invention relates to a method for determining cancer susceptibility
CC   in a human patient. The method involves determining the identity of the
CC   nucleotide at position 457 of the serine/threonine kinase 15 (STK15) DNA
CC   (ABZ75005). This site is a T/A single nucleotide polymorphism (SNP) in
CC   the coding region of the DNA, resulting in either a Phe or Ile residue at
CC   position 31 in the corresponding STK15 protein (ABP97366). The A457
CC   (Ile31) allele (see ABZ75006, ABP97367) is associated with an increased
CC   cancer susceptibility. STK15 (also known as STK6 and Aurora2) is a
CC   centrosome-associated kinase that is highly expressed at the G2 and M
CC   phase of the cell cycle, and its gene is located on chromosome 20. The
CC   method of the invention are useful for determining cancer susceptibility
CC   and for prognosing, detecting and/or diagnosing cancers such as malignant
CC   astrocytoma, glioblastoma, medulloblastoma, gastric cancer, colorectal
CC   cancer, colorectal adenoma, acute myelogenous leukemia, lung cancer,
CC   renal cancer, leukaemia, breast cancer, prostate cancer, endometrial
CC   cancer and neuroblastoma. This sequence represents the Ile31 variant of
CC   human STK15 which is associated with increased cancer susceptibility.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the human STK15 sequence shown in the sequence listing
XX
SQ   Sequence 403 AA;

  Query Match             99.7%;  Score 2104;  DB 6;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 3.5e-186;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 10
ABR48160
ID   ABR48160 standard; protein; 403 AA.
XX
AC   ABR48160;
XX
DT   12-JUN-2003  (first entry)
XX
DE   Human bladder cancer associated protein sequence SEQ ID NO:35.
XX
KW   Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003003906-A2.
XX
PD   16-JAN-2003.
XX
PF   03-JUL-2002; 2002WO-US021338.
XX
PR   03-JUL-2001; 2001US-0302814P.
PR   03-AUG-2001; 2001US-0310099P.
PR   08-NOV-2001; 2001US-0343705P.
PR   13-NOV-2001; 2001US-0350666P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Mack DH,  Aziz N;
XX
DR   WPI; 2003-201532/19.
DR   N-PSDB; ACC50970.
XX
PT   Detecting a bladder cancer-associated transcript in a cell from a
PT   patient, comprises contacting a biological sample from the patient with a
PT   bladder cancer-associated polynucleotide or antibody.
XX
PS   Claim 10; Page 240-241; 307pp; English.
XX
CC   The present invention describes a method for detecting a bladder cancer-
CC   associated transcript in a cell from a patient. The method comprises
CC   contacting a biological sample from the patient with a polynucleotide
CC   that selectively hybridises to a sequence that is 80 % identical to a
CC   table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC   encode the human bladder cancer-associated proteins given in ABR48146 to
CC   ABR48242). Bladder cancer-associated sequences from the present invention
CC   have cytostatic activities, and can be used in antisense gene therapy and
CC   in vaccine production. The method can be used for detecting a bladder
CC   cancer-associated transcript in a cell from a patient. The method is
CC   useful in diagnosing or treating bladder cancer and in screening for
CC   compounds that modulate bladder cancer, such as hormones or antibodies.
CC   The nucleic acid molecules from the present invention may be used in
CC   various screening and diagnostic methods, and for gene therapy, vaccine
CC   and/or antisense/inhibition applications
XX
SQ   Sequence 403 AA;

  Query Match             99.7%;  Score 2104;  DB 6;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 3.5e-186;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
```

```
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 11
ADN39888
ID   ADN39888 standard; protein; 403 AA.
XX
AC   ADN39888;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C258.
XX
KW   Human; differential expression; cancer; angiogenic disorder;
KW   fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW   inflammatory disease; autoimmune disease;
KW   retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW   detection; diagnosis; prognosis; drug screening; drug targeting;
KW   wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW   vulnerary; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003042661-A2.
XX
PD   22-MAY-2003.
XX
PF   13-NOV-2002; 2002WO-US036810.
XX
PR   13-NOV-2001; 2001US-0350666P.
PR   21-NOV-2001; 2001US-0332464P.
PR   29-NOV-2001; 2001US-0334393P.
PR   03-DEC-2001; 2001US-0335394P.
PR   14-DEC-2001; 2001US-0340376P.
PR   08-JAN-2002; 2002US-0347211P.
PR   10-JAN-2002; 2002US-0347349P.
PR   08-FEB-2002; 2002US-0355250P.
PR   13-FEB-2002; 2002US-0356714P.
PR   20-FEB-2002; 2002US-0359077P.
PR   29-MAR-2002; 2002US-0368809P.
PR   04-APR-2002; 2002US-0370110P.
PR   12-APR-2002; 2002US-0372246P.
PR   05-JUN-2002; 2002US-0386614P.
PR   16-JUL-2002; 2002US-0396839P.
PR   22-JUL-2002; 2002US-0397775P.
PR   22-JUL-2002; 2002US-0397845P.
PR   09-SEP-2002; 2002US-0409450P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D,  Aziz N,  Ginsburg WM,  Gish KC,  Glynne R,  Hevezi PA;
PI   Mack DH,  Murray R,  Watson SR,  Wilson KE,  Zlotnik A;
XX
DR   WPI; 2003-468649/44.
DR   N-PSDB; ADN39671.
XX
PT   Determining the presence or absence of a pathological cell in a patient,
PT   useful for diagnosing, prognosing or treating cancer, comprises detecting
PT   a nucleic acid in a biological sample.
XX
PS   Claim 12; SEQ ID NO C258; 1385pp; English.
XX
CC   The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC   whose expression is upregulated or downregulated in specific cancers or
CC   other diseases such as angiogenic or fibrotic disorders, and to methods
CC   of determining the presence or absence of a pathological cell in a
CC   patient by detecting a nucleic acid at least 80% identical to those of
CC   the invention or by detecting a polypeptide of the invention. The
CC   invention also relates to expression vectors and host cells comprising a
CC   nucleic acid of the invention; antibodies which specifically bind a
CC   polypeptide of the invention; use of such antibodies for drug targeting;
CC   and methods of screening for modulators of activity or expression of the
CC   polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC   antibodies and methods are useful for diagnosing, prognosing and treating
CC   cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC   atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC   neovascularistaion syndromes, scarring and uterine fibroids. They may
CC   also be useful in wound healing and in contraception. The present
CC   sequence represents a polypeptide of the invention.
XX
SQ   Sequence 403 AA;

  Query Match             99.7%;  Score 2104;  DB 7;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 3.5e-186;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 12
ADE15337
ID   ADE15337 standard; protein; 403 AA.
XX
AC   ADE15337;
XX
DT   29-JAN-2004  (first entry)
```

XX
DE Truncated Aurora-2 kinase #SEQ ID 1.
XX
KW Aurora-2 kinase; three dimensional structure; protein co-ordinate data.
XX
OS Homo sapiens.
XX
PN WO2003092607-A2.
XX
PD 13-NOV-2003.
XX
PF 01-MAY-2003; 2003WO-US013605.
XX
PR 01-MAY-2002; 2002US-0377510P.
XX
PA (VERT-) VERTEX PHARM INC.
XX
PI Cheetham G, Knegtel R, Swenson L, Coll JT, Renwick S, Weber P;
XX
DR WPI; 2004-022617/02.
XX
PT Crystal useful for screening, designing and evaluating compounds as
PT agonists or antagonists of Aurora-2 kinases, comprises optionally
PT phosphorylated Aurora-2 kinase domain.
XX
PS Example 1; SEQ ID NO 1; 242pp; English.
XX
CC The invention relates to a crystal comprising an optionally
CC phosphorylated Aurora-2 kinase domain or its homologue. The protein of
CC the invention may be used for identifying agonists or antagonists of
CC Aurora-2 kinases, and for generation of three dimensional structures and
CC structural co-ordinate information of Aurora family proteins. The
CC crystalline structure facilitates the designing of selective inhibitors
CC of Aurora family kinases (particularly Aurora-2 kinase) so that
CC undesirable side effects associated with non-selective inhibitors can be
CC avoided. The structural co-ordinates solve the structure of Aurora-2
CC proteins that have amino acid substitutions, additions and/or mutations,
CC and serve as additional tools to determine the most efficient binding
CC interactions. The current sequence represents the Aurora-2 kinase amino
CC acid sequence.
XX
SQ Sequence 403 AA;

Query Match 99.7%; Score 2104; DB 8; Length 403;
Best Local Similarity 99.8%; Pred. No. 3.5e-186;
Matches 402; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 13
ADH59556
ID ADH59556 standard; protein; 403 AA.
XX
AC ADH59556;
XX
DT 25-MAR-2004 (first entry)
XX
DE Monoclonal antibody of the invention.
XX
KW monoclonal antibody; MAb; aurora-A kinase; Cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003106500-A1.
XX
PD 24-DEC-2003.
XX
PF 12-JUN-2003; 2003WO-FR001772.
XX
PR 12-JUN-2002; 2002FR-00007212.
XX
PA (CNRS) CENT NAT RECH SCI.
PA (FRSA-) ETAB FR DU SANG-BRETAGNE.
XX
PI Prigent C, Martin A;
XX
DR WPI; 2004-071550/07.
DR N-PSDB; ADH59555.
XX
PT New monoclonal antibody specific for aurora-A kinase, useful for
PT diagnosis, prognosis and treatment of solid tumors, also for drug
PT screening.
XX
PS Claim 8; SEQ ID NO 2; 38pp; French.
XX
CC The present invention relates to monoclonal antibody (MAb) that
CC recognizes specifically the human or murine aurora-A kinase binds to
CC membranes that contain, can detect and optionally purify by
CC immunoprecipitation, stains biological tissues where is secreted and does
CC not inhibit the enzymatic activity of MAb. Cytostatic. MAb are used for
CC in vitro diagnosis and prognosis of cancers in humans and animals,
CC particularly breast, gastric and colorectal cancer; for treatment of
CC these cancers; and to screen for inhibitors of MAb. The present sequence
CC represents the monoclonal antibody of the invention.
XX
SQ Sequence 403 AA;

Query Match 99.7%; Score 2104; DB 8; Length 403;
Best Local Similarity 99.8%; Pred. No. 3.5e-186;
Matches 402; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 14
ADK67720
ID ADK67720 standard; protein; 403 AA.
XX
AC ADK67720;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human modifier of AXIN (MAX) polypeptide.
XX
KW Human; modifier of AXIN; MAX; cytostatic; gene therapy; protein kinase;
KW enzyme.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT Misc-difference 31
FT   /note= "Encoded by TTT"
FT Misc-difference 104
FT   /note= "Encoded by TC"
FT Misc-difference 128
FT   /note= "Encoded by TG"
FT Misc-difference 235
FT   /note= "Encoded by AACT"
FT Misc-difference 240
FT   /note= "Encoded by TT"
XX
PN WO2004013308-A2.
XX
PD 12-FEB-2004.
XX
PF 06-AUG-2003; 2003WO-US024560.
XX
PR 06-AUG-2002; 2002US-0401534P.
PR 16-SEP-2002; 2002US-0411153P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Gendreau SB, Dora EG, Lickteig K, Amundsen CD;
XX
DR WPI; 2004-157122/15.
DR N-PSDB; ADK67713.
XX
PT Identifying a candidate AXIN pathway-modulating agent, useful in
PT diagnosing and treating cancer, comprises providing an assay system
PT comprising a MAX polypeptide or nucleic acid.
XX
PS Example 1; SEQ ID NO 11; 74pp; English.
XX
CC The present sequence is that of a human modifier of AXIN (MAX)
CC polypeptide characterised as a serine/threonine kinase. Genetic screens
CC were designed to identify modifiers of the axin pathway in Caenorhabditis
CC elegans, where a reduction of function pry-1 (axin) mutant was used.
CC Genes causing altered phenotypes were isolated, and their human
CC homologues, termed MAX, were identified, including the present sequence.
CC These MAX polynucleotides and polypeptides are attractive targets for the
CC treatment of pathologies associated with a defective AXIN signalling
CC pathway, such as cancer. Modulation of MAX or its binding partner is
CC useful for understanding the association of the AXIN pathway and its
CC members in normal and disease conditions and for developing diagnostic
CC and therapeutic modalities for AXIN related pathologies. MAX-modulating
CC agents that act by inhibiting or enhancing MAX expression, directly or
CC indirectly, e.g. by affecting MAX function can be identified using
CC methods of the invention. MAX modulating agents are useful in diagnosis,
CC therapy and pharmaceutical development. Preferred MAX modulating agents
CC include antisense and phosphorothioate morpholino oligomers.
XX
SQ Sequence 403 AA;

```
  Query Match             99.7%;  Score 2104;  DB 8;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 3.5e-186;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 15
ADM72205
ID ADM72205 standard; protein; 403 AA.
XX
AC ADM72205;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human TASK104 polypeptide.
XX
KW TASK; tumour-associated kinase; cytostatic; tumour antigen;
KW cell proliferative disorder; cancer; transgenic; human.
XX
OS Homo sapiens.
XX
PN WO2004024064-A2.
XX
PD 25-MAR-2004.
XX
PF 05-SEP-2003; 2003WO-US027894.
XX
PR 11-SEP-2002; 2002US-0410166P.
XX
PA (GETH) GENENTECH INC.
XX
PI Desauvage FJ, Wood WI, Zhang Z;
XX
DR WPI; 2004-282985/26.
DR N-PSDB; ADM72204.
XX
PT New tumor-associated kinase nucleic acids and polypeptides, useful as
PT hybridization probes for isolating full length TASK DNA, for generating
PT transgenic animals, in chromosome identification, or for tissue typing.

```
XX
PS   Claim 12; SEQ ID NO 10; 163pp; English.
XX
CC   The invention relates to new isolated tumour-associated kinase (TASK)
CC   nucleic acid molecules and encoded polypeptides. Cytostatic. The
CC   antibody, oligopeptide or organic molecule that binds to the TASK
CC   polypeptide are useful for treating a mammal having a tumour comprising
CC   cells expressing the polypeptide. Antagonists of TASK are useful for
CC   treating or preventing a cell proliferative disorder (e.g. cancer)
CC   associated with increased expression or activity of TASK polypeptide. The
CC   TASK polynucleotides and polypeptides may be used as hybridization probes
CC   for isolating full length TASK DNA, for generating transgenic animals, in
CC   chromosome identification, or for tissue typing. The present sequence
CC   represents a human TASK polypeptide.
XX
SQ   Sequence 403 AA;

  Query Match             99.7%;  Score 2104;  DB 8;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 3.5e-186;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQIPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 16
ABP97366
ID   ABP97366 standard; protein; 403 AA.
XX
AC   ABP97366;
XX
DT   10-MAY-2003  (first entry)
XX
DE   Human serine/threonine kinase 15 (STK15).
XX
KW   Human; serine/threonine kinase 15; STK15; STK6; Aurora2; cell cycle;
KW   chromosome 20; centrosome-associated kinase; cancer susceptibility;
KW   single nucleotide polymorphism; SNP; genetic diagnosis; prognosis;
KW   detection; diagnosis; cancer; malignant astrocytoma; glioblastoma;
KW   medulloblastoma; gastric cancer; colorectal cancer; colorectal adenoma;
KW   acute myelogenous leukaemia; lung cancer; renal cancer; leukaemia;
KW   breast cancer; prostate cancer; endometrial cancer; neuroblastoma;
KW   enzyme.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 31
FT                   /label= Phe, Ile
FT                   /note= "Encoded by WTT. This position represents a
FT                   polymorphic site"
FT   Misc-difference 57
FT                   /note= "Encoded by ATT"
XX
PN   WO2003012046-A2.
XX
PD   13-FEB-2003.
XX
PF   29-JUL-2002; 2002WO-US024115.
XX
PR   27-JUL-2001; 2001US-0308911P.
PR   28-NOV-2001; 2001US-0334146P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Toland AE,  Balmain A;
XX
DR   WPI; 2003-239517/23.
DR   N-PSDB; ABZ75005.
XX
PT   Determining cancer susceptibility in a human subject comprises
PT   identifying in a nucleic acid sample from the subject, a nucleotide
PT   occurrence of a single polynucleotide polymorphism (SNP) of the STK15
PT   gene.
XX
PS   Claim 63; Page 85-86; 92pp; English.
XX
CC   The invention relates to a method for determining cancer susceptibility
CC   in a human patient. The method involves determining the identity of the
CC   nucleotide at position 457 of the serine/threonine kinase 15 (STK15) DNA
CC   (ABZ75005). This site is a T/A single nucleotide polymorphism (SNP) in
CC   the coding region of the DNA, resulting in either a Phe or Ile residue at
CC   position 31 in the corresponding STK15 protein (ABP97366). The A457
CC   (Ile31) allele (see ABZ75006, ABP97367) is associated with an increased
CC   cancer susceptibility. STK15 (also known as STK6 and Aurora2) is a
CC   centrosome-associated kinase that is highly expressed at the G2 and M
CC   phase of the cell cycle, and its gene is located on chromosome 20. The
CC   method of the invention are useful for determining cancer susceptibility
CC   and for prognosing, detecting and/or diagnosing cancers such as malignant
CC   astrocytoma, glioblastoma, medulloblastoma, gastric cancer, colorectal
CC   cancer, colorectal adenoma, acute myelogenous leukemia, lung cancer,
CC   renal cancer, leukaemia, breast cancer, prostate cancer, endometrial
CC   cancer and neuroblastoma. This sequence represents human STK15
XX
SQ   Sequence 403 AA;

  Query Match             99.7%;  Score 2103;  DB 6;  Length 403;
  Best Local Similarity   99.8%;  Pred. No. 4.3e-186;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQXPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
                ||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 17
ADJ31727
ID   ADJ31727 standard; protein; 403 AA.
XX
AC   ADJ31727;
XX
DT   22-APR-2004  (first entry)
XX
DE   Human mitotic kinase hARK.
XX
KW   Human; enzyme; SAK; Snk Akin Kinase; serine/threonine protein kinase;
KW   mitotic kinase; cellular proliferation; chemosensitivity; cancer;
KW   proliferative disorder; cytostatic; hARK.
XX
OS   Homo sapiens.
XX
PN   US2003027756-A1.
XX
PD   06-FEB-2003.
XX
PF   21-DEC-2001; 2001US-00026021.
XX
PR   01-AUG-2001; 2001US-0309632P.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Hitoshi Y,  Demo S,  Jenkins Y;
XX
DR   WPI; 2003-479546/45.
XX
PT   Modulating cellular proliferation in subject, by administering
PT   serine/threonine protein kinase involved in modulation of cellular
PT   proliferation and cell cycle regulation, or a nucleic acid encoding the
PT   polypeptide.
XX
PS   Example 1; Fig 2; 41pp; English.
XX
CC   The invention relates to modulating (M1) cellular proliferation in a
CC   subject, involves administering to the subject a serine/threonine protein
CC   kinase involved in modulation of cellular proliferation and cell cycle
CC   regulation (SAK polypeptide, Snk Akin Kinase, a mitotic kinase) encoded
CC   by a nucleic acid that hybridises under stringent conditions to a nucleic
CC   acid encoding a polypeptide appearing as ADJ31723. Also included are
CC   identifying (M2) a compound capable of interfering with binding of SAK or
CC   its fragment (by combining SAK or its fragment with Chk2 polypeptide and
CC   the compound, where SAK or its fragment has kinase activity, and
CC   determining the binding of SAK or its fragment to Chk2), identifying (M3)
CC   a compound that modulates cellular proliferation (by contacting the
CC   compound with SAK, and determining the functional effect of the compound
CC   on SAK), identifying (M4) a compound that modulates cellular
CC   proliferation or chemosensitivity, (by contacting the compound with SAK
CC   or its fragment, and determining the physical effect of the compound on
CC   SAK), determining the chemical or phenotypic effect of the compound upon
CC   a cell comprising SAK or its fragment and modulating (M5) cellular
CC   proliferation in a subject (by administering to the subject a compound
CC   identified by M3). M1 or M5 is useful for modulating cellular
CC   proliferation in a subject, preferably a human having cancer. M2 is
CC   useful for identifying a compound capable of interfering with binding of
CC   SAK or its fragment. M3 is useful for identifying a compound that
CC   modulates cellular proliferation. M4 is useful for identifying a compound
CC   that modulates cellular proliferation or chemosensitivity, where the
CC   compound is an antibody, antisense molecule, small organic molecule, or
CC   circular peptide. M1 or M5 or the compound identified by M3 or M4 is
CC   useful for treating cancer and other proliferative disorders. The present
CC   sequence represents a human mitotic kinase similar to SAK.
XX
SQ   Sequence 403 AA;

  Query Match             99.3%;  Score 2095;  DB 7;  Length 403;
  Best Local Similarity   99.3%;  Pred. No. 2.4e-185;
  Matches  400;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQQTSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              ||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403


RESULT 18
ABR61579
ID   ABR61579 standard; protein; 402 AA.
XX
AC   ABR61579;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human HsAIRK-1 (Aurora B) protein.
XX
KW   RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; HsAIRK-1;
KW   Aurora B.
XX
OS   Homo sapiens.
XX
PN   WO2003087395-A2.
XX
PD   23-OCT-2003.
XX
PF   15-APR-2003; 2003WO-IB002972.
XX
PR   15-APR-2002; 2002US-0372483P.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Garbay C,  Gigoux V,  Camonis J,  L'hoste S,  Samson J;
XX
DR   WPI; 2003-845340/78.
DR   N-PSDB; ACF58068.
XX
PT   Identifying anti-tumoral compounds, comprises determining the capacity of
PT   a compound to inhibit interaction between RasGAP and Drosophila
PT   melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and
PT   aurora binding protein.
XX
PS   Disclosure; Page 59-60; 0pp; English.
XX
```

```
CC   The invention relates to identifying a biologically active compound with
CC   anti-tumoural properties, where the compound is studied for its capacity
CC   to inhibit the interaction between: (a) RasGAP and the Drosophila
CC   melanogaster Aurora kinase or an orthologue of the kinase, or its
CC   fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding
CC   protein and the D. melanogaster Aurora kinase or an orthologue of the
CC   kinase, or its fragment. The method is useful for identifying anti-
CC   tumoural agents. The present sequence represents a human HsAIRK-1 (Aurora
CC   B) protein, an orthologue of the D. melanogaster Aurora kinase
XX
SQ   Sequence 402 AA;

  Query Match             91.7%;  Score 1935.5;  DB 7;  Length 402;
  Best Local Similarity   92.9%;  Pred. No. 1.5e-170;
  Matches  378;  Conservative    3;  Mismatches   17;  Indels    9;  Gaps    3;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPS----APENNPEEELAS 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  |  :
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSHLKIILRRNWHQN--R 118

Qy        117 KQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEH 176
              | ||:   |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 KMKNQ---KEAVALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEH 175

Qy        177 QLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATY 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 QLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTANL 235

Qy        237 ITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYL 296
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 YNRIANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYL 295

Qy        297 PPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGA 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 PPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGA 355

Qy        357 RDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        356 RDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 402


RESULT 19
ADD89969
ID   ADD89969 standard; protein; 402 AA.
XX
AC   ADD89969;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human cancer-associated protein kinase STK6.
XX
KW   STK6; protein kinase; enzyme; cancer; cytostatic; immunosuppressive;
KW   antidiabetic; antirheumatic; antiarthritic; antipsoriatic;
KW   antiangiogenic; antiarteriosclerotic; antiinflammatory; vulnerary;
KW   gynaecological; neuroprotective; gene therapy; human.
XX
OS   Homo sapiens.
XX
PN   WO2003083096-A2.
XX
PD   09-OCT-2003.
XX
PF   21-MAR-2003; 2003WO-CA000409.
XX
PR   28-MAR-2002; 2002US-0368853P.
XX
PA   (KINE-) KINETEK PHARM INC.
XX
PI   Delaney AD;
XX
DR   WPI; 2003-833542/77.
DR   N-PSDB; ADD89968.
XX
PT   New nucleic acids encoding cancer associated protein kinases, useful as
PT   targets for screening pharmaceutical agents that inhibit the growth of
PT   tumor cells, or for diagnosing and treating cancer, inflammation or
PT   autoimmune disease.
XX
PS   Claim 1; Page 73-74; 91pp; English.
XX
CC   The present sequence is the protein sequence of the human cancer-
CC   associated protein kinase, STK6. Serine/threonine kinase 6 (STK6),
CC   otherwise known as AIK, is an aurora/IPL1-like kinase possibly involved
CC   in centrosome function. STK6 gene expression is significantly up-
CC   regulated in cancers of the liver, muscle, placenta, and prostate. STK6
CC   is one of a set of protein kinases that are shown by the invention to be
CC   over-expressed in hyper-proliferative cells. These protein kinases
CC   provide targets for drug screening for agents effective in inhibiting the
CC   growth or metastasis of tumour cells, and for determining other molecular
CC   targets in kinase signal transduction pathways involved in transformation
CC   and growth of tumour cells. A claimed method for inhibiting the growth of
CC   a cancer cell involves down-regulating the activity of the protein kinase
CC   using an antisense sequence or inhibitor of kinase activity, especially
CC   where the cancer cell is a breast, liver, colon, muscle, prostate,
CC   kidney, lung, placental or uterine cancer cell. Detection of over-
CC   expression in cancers provides a useful diagnostic for predicting patient
CC   prognosis and probability of drug effectiveness. Agents that specifically
CC   bind the protein kinases can be used for treatment and visualisation of
CC   tumours in patients. The protein kinase polypeptides and nucleic acids
CC   may also be used for treating hyperproliferative diseases, such as
CC   autoimmune disease, diabetes mellitus, multiple sclerosis, rheumatoid
CC   arthritis, psoriasis, atherosclerosis, inflammation, scarring,
CC   endometriosis and angiogenesis.
XX
SQ   Sequence 402 AA;

  Query Match             91.7%;  Score 1935.5;  DB 7;  Length 402;
  Best Local Similarity   92.9%;  Pred. No. 1.5e-170;
  Matches  378;  Conservative    3;  Mismatches   17;  Indels    9;  Gaps    3;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPS----APENNPEEELAS 116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||  |  :
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSHLKIILRRNWHQN--R 118

Qy        117 KQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEH 176
              | ||:   |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 KMKNQ---KEAVALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEH 175

Qy        177 QLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATY 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 QLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTANL 235

Qy        237 ITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYL 296
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 YNRIANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYL 295

Qy        297 PPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGA 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 PPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGA 355

Qy        357 RDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        356 RDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 402


RESULT 20
AAO18740
ID   AAO18740 standard; protein; 403 AA.
```

```
XX
AC   AAO18740;
XX
DT   24-OCT-2002  (first entry)
XX
DE   Human NOV4 protein.
XX
KW   Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
KW   storage disorder; muscle disorder; neurodegenerative disorder; nootropic;
KW   developmental defect; neuroprotective; antiparkinsonian; hypotensive;
KW   hypertensive; haemostatic; cardiant; antianginal; dermatological;
KW   immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
KW   antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;
KW   vulnerary; anorectic; antidiabetic; immunomodulator; antipsoriatic;
KW   nephrotropic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
KW   antiinfertility; antimanic; antidepressant; metabolic; cytostatic;
KW   tranquilizer; analgesic.
XX
OS   Homo sapiens.
XX
PN   WO200257450-A2.
XX
PD   25-JUL-2002.
XX
PF   29-NOV-2001; 2001WO-US048922.
XX
PR   29-NOV-2000; 2000US-0253834P.
PR   30-NOV-2000; 2000US-0250926P.
PR   25-JAN-2001; 2001US-0264180P.
PR   20-AUG-2001; 2001US-0313656P.
PR   05-OCT-2001; 2001US-0327456P.
PR   28-NOV-2001; 2001US-00327456.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Edinger S,  Macdougall JR,  Millet I,  Ellerman K,  Stone DJ;
PI   Gerlach V,  Grosse WM,  Alsobrook JP,  Lepley DM,  Rieger D,  Burgess CE;
PI   Casman SJ,  Spytek KA,  Boldog FL,  Li L,  Padigaru M,  Mishra V;
PI   Patturajan M,  Shenoy S,  Rastelli L,  Tchernev VT,  Vernet CAM;
PI   Zerhusen BD,  Malyankar UM,  Guo X,  Miller CE,  Gangolli EA;
XX
DR   WPI; 2002-590741/63.
DR   N-PSDB; ABT06285.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing in NOVX-associated disorders e.g. cardiomyopathy,
PT   atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
XX
PS   Claim 1; Page 62; 353pp; English.
XX
CC   The present invention provides the protein and coding sequences of
CC   several novel human proteins, designated NOVX. These can be used in the
CC   treatment of, amongst others, cancers, autoimmune diseases, infections,
CC   inflammatory diseases, storage disorders, muscle disorders,
CC   neurodegenerative diseases and developmental defects. The present
CC   sequence is a protein of the invention
XX
SQ   Sequence 403 AA;

  Query Match             91.2%;  Score 1924;  DB 5;  Length 403;
  Best Local Similarity   91.8%;  Pred. No. 1.8e-169;
  Matches  370;  Conservative   11;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||:|:||||||||:|  |||||: ||||||||||||| |||||| ||||||||||||||
Db          1 MDQSEENCISGPVEAKTPVGGPEHVLVTQQFPCQNPLPANSGQAQWVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||||||||||||||||||||||||| |||||||| ||| ||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPASRPLNNTQNSKQSPLSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |||||||||||| ||||| ||||||||||||||||||||:||||||||||||||||||
Db        121 EESKKRQWALEDLEIGRPPGKGKFGNVYLAREKQSKFILALRVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||:||||:||||||||||||||||| || ||  ||||||||||||||||||||
Db        181 EVEIQSHLQHPNIIRLYGYFHDATRVYLILEYTPLETVNTELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              |:|||||||| ||||||||||||||||||:||:|||| ||||||||||||||||||||||
Db        241 ASALSYCHSKTVIHRDIKPENLLLGSAGELEIANFGWSEHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCCEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||| ||||||||||:|||||||||||||||||| || |
Db        361 SRLLKHVPSQRPMLREVLEYPWITANSSKPSNCQNKESTSKYS 403


RESULT 21
ABP97470
ID   ABP97470 standard; protein; 319 AA.
XX
AC   ABP97470;
XX
DT   19-SEP-2003  (first entry)
XX
DE   Truncated mutant human Aurora A kinase, [T287] Aurora A (94-402).
XX
KW   Aurora A kinase; human; catalytic domain; chromosome 20q13; 3D structure;
KW   X-ray crystallography; protein co-ordinate data; serine/threonine kinase;
KW   cell cycle regulation; G2/M boundary; oncogenesis;
KW   hyperproliferative disorder; cancer; tumour; breast; colon; colorectal;
KW   Aurora homologue design; rational drug design; drug screening; enzyme;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .10
FT                   /note= "pTB375NBSE vector-encoded sequence, comprising a
FT                   hexahistidine sequence"
FT   Domain          42. .125
FT                   /label= N-terminal_domain
FT                   /note= "Corresponds to residues 125-208 of the native
FT                   human Aurora A kinase"
FT   Active-site     126. .131
FT                   /label= ATP_binding_pocket
FT                   /note= "Corresponds to residues 209-214 of the native
FT                   human Aurora A kinase. Also referred to as a hinge region
FT                   between the N- and C-terminal domains"
FT   Binding-site    127
FT                   /note= "The peptide carbonyl group of this residue binds
FT                   adenine nitrogen atom N6. Corresponds to residue 210 of
FT                   the native human Aurora A kinase"
FT   Binding-site    129
FT                   /note= "The main chain nitrogen of this residue binds
FT                   adenine nitrogen atom N1. Corresponds to residue 212 of
FT                   the native human Aurora A kinase"
FT   Domain          132. .291
FT                   /label= C-terminal_domain
FT                   /note= "Corresponds to residues 215-374 of the native
FT                   human Aurora A kinase"
FT   Misc-difference 204
FT                   /note= "Asp replaces wild-type Thr. Corresponds to
FT                   position 287 of the native human Aurora A kinase"
XX
PN   WO2003031606-A2.
XX
PD   17-APR-2003.
XX
PF   08-OCT-2002; 2002WO-GB004589.
```

```
XX
PR   10-OCT-2001; 2001GB-00024299.
XX
PA   (ASTR ) ASTRAZENECA AB.
PA   (ASTR ) ASTRAZENECA UK LTD.
XX
PI   Anderson M,  Keen NJ,  Pannifer ADB,  Pauptit RA,  Rowsell S;
XX
DR   WPI; 2003-403104/38.
XX
PT   Novel crystalline form of polypeptide comprising catalytic domain of
PT   Aurora kinase, whose atomic coordinate data is useful for designing or
PT   selecting Aurora chemical inhibitor, and for designing an Aurora protein.
XX
PS   Example 1; Page 117; 137pp; English.
XX
CC   The invention relates to the three-dimensional structure of human Aurora
CC   A kinase. The invention describes the methods used to prepare and
CC   crystallise truncated mutant forms of human Aurora A kinase comprising
CC   the catalytic domain, which were then subjected to X-ray crystallography
CC   to generate atomic coordinates. The mutant forms of Aurora A used in the
CC   invention all contain a T287D substitution, which produces a mimic of the
CC   activated protein which can be provided as a homogeneous sample (native
CC   Aurora A kinase is activated by phosphorylation of T287). Two crystalline
CC   forms of the Aurora A kinase catalytic domain were produced - one
CC   complexed with the hydrolysable ATP analogue ANP-PNP, and one complexed
CC   with a synthetic inhibitor. The ATP analogue was found to occupy a cleft
CC   between the N-terminal domain (residues 125-208) and the C-terminal
CC   domain (residues 215-374). Aurora A kinase is a serine/threonine kinase
CC   which is involved in cell cycle regulation, with its expression and
CC   activity peaking at the G2/M boundary. It has been implicated in
CC   oncogenesis - the Aurora A gene maps to chromosome 20q13, a region which
CC   is frequently amplified in human tumours including breast and colon
CC   tumours, and it is overexpressed in over 50% of primary colorectal
CC   tumours. The crystal structure of human Aurora A kinase may be used to
CC   select or design chemical modulators of Aurora kinase, particularly
CC   Aurora kinase inhibitors. These modulators may be used to prevent or
CC   treat undesirable physical and pharmacological consequences of
CC   inappropriate Aurora activity, especially cancers or other
CC   hyperproliferative disorders. The Aurora A kinase 3D structure may also
CC   be used in designing an Aurora protein or Aurora homologue, and in
CC   elucidating the 3D structure of other proteins with structural similarity
CC   to Aurora kinases. The present sequence represents a truncated T287D
CC   mutant Aurora A kinase sequence which was prepared and crystallised in an
CC   example from the invention. This sequence comprises the Aurora A kinase
CC   domain and also contains a vector-encoded sequence at the N-terminus
XX
SQ   Sequence 319 AA;

  Query Match             76.5%;  Score 1614;  DB 6;  Length 319;
  Best Local Similarity   99.4%;  Pred. No. 7.4e-141;
  Matches  308;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          94 NTQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREK 153
               :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          10 STQKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREK 69

Qy         154 QSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYA 213
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          70 QSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYA 129

Qy         214 PLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIA 273
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         130 PLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIA 189

Qy         274 DFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANT 333
               ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         190 DFGWSVHAPSSRRTDLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANT 249

Qy         334 YQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNC 393
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         250 YQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNC 309

Qy         394 QNKESASKQS 403
               ||||||||||
Db         310 QNKESASKQS 319


RESULT 22
ABP97472
ID   ABP97472 standard; protein; 309 AA.
XX
AC   ABP97472;
XX
DT   19-SEP-2003  (first entry)
XX
DE   Truncated mutant human Aurora A kinase, [T287] Aurora A (113-400).
XX
KW   Aurora A kinase; human; catalytic domain; chromosome 20q13; 3D structure;
KW   X-ray crystallography; protein co-ordinate data; serine/threonine kinase;
KW   cell cycle regulation; G2/M boundary; oncogenesis;
KW   hyperproliferative disorder; cancer; tumour; breast; colon; colorectal;
KW   Aurora homologue design; rational drug design; drug screening; enzyme;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .21
FT                   /note= "pET28a vector-encoded sequence, comprising a
FT                   hexahistidine sequence"
FT   Domain          33. .117
FT                   /label= N-terminal_domain
FT                   /note= "Corresponds to residues 125-208 of the native
FT                   human Aurora A kinase"
FT   Active-site     118. .123
FT                   /label= ATP_binding_pocket
FT                   /note= "Corresponds to residues 209-214 of the native
FT                   human Aurora A kinase. Also referred to as a hinge region
FT                   between the N- and C-terminal domains"
FT   Binding-site    119
FT                   /note= "The peptide carbonyl group of this residue binds
FT                   adenine nitrogen atom N6. Corresponds to residue 210 of
FT                   the native human Aurora A kinase"
FT   Binding-site    121
FT                   /note= "The main chain nitrogen of this residue binds
FT                   adenine nitrogen atom N1. Corresponds to residue 212 of
FT                   the native human Aurora A kinase"
FT   Domain          124. .283
FT                   /label= C-terminal_domain
FT                   /note= "Corresponds to residues 215-374 of the native
FT                   human Aurora A kinase"
FT   Misc-difference 196
FT                   /note= "Asp replaces wild-type Thr. Corresponds to
FT                   position 287 of the native human Aurora A kinase"
XX
PN   WO2003031606-A2.
XX
PD   17-APR-2003.
XX
PF   08-OCT-2002; 2002WO-GB004589.
XX
PR   10-OCT-2001; 2001GB-00024299.
XX
PA   (ASTR ) ASTRAZENECA AB.
PA   (ASTR ) ASTRAZENECA UK LTD.
XX
PI   Anderson M,  Keen NJ,  Pannifer ADB,  Pauptit RA,  Rowsell S;
XX
DR   WPI; 2003-403104/38.
XX
PT   Novel crystalline form of polypeptide comprising catalytic domain of
PT   Aurora kinase, whose atomic coordinate data is useful for designing or
PT   selecting Aurora chemical inhibitor, and for designing an Aurora protein.
XX
```

```
PS   Example 1; Page 119; 137pp; English.
XX
CC   The invention relates to the three-dimensional structure of human Aurora
CC   A kinase. The invention describes the methods used to prepare and
CC   crystallise truncated mutant forms of human Aurora A kinase comprising
CC   the catalytic domain, which were then subjected to X-ray crystallography
CC   to generate atomic coordinates. The mutant forms of Aurora A used in the
CC   invention all contain a T287D substitution, which produces a mimic of the
CC   activated protein which can be provided as a homogeneous sample (native
CC   Aurora A kinase is activated by phosphorylation of T287). Two crystalline
CC   forms of the Aurora A kinase catalytic domain were produced - one
CC   complexed with the hydrolysable ATP analogue ANP-PNP, and one complexed
CC   with a synthetic inhibitor. The ATP analogue was found to occupy a cleft
CC   between the N-terminal domain (residues 125-208) and the C-terminal
CC   domain (residues 215-374). Aurora A kinase is a serine/threonine kinase
CC   which is involved in cell cycle regulation, with its expression and
CC   activity peaking at the G2/M boundary. It has been implicated in
CC   oncogenesis - the Aurora A gene maps to chromosome 20q13, a region which
CC   is frequently amplified in human tumours including breast and colon
CC   tumours, and it is overexpressed in over 50% of primary colorectal
CC   tumours. The crystal structure of human Aurora A kinase may be used to
CC   select or design chemical modulators of Aurora kinase, particularly
CC   Aurora kinase inhibitors. These modulators may be used to prevent or
CC   treat undesirable physical and pharmacological consequences of
CC   inappropriate Aurora activity, especially cancers or other
CC   hyperproliferative disorders. The Aurora A kinase 3D structure may also
CC   be used in designing an Aurora protein or Aurora homologue, and in
CC   elucidating the 3D structure of other proteins with structural similarity
CC   to Aurora kinases. The present sequence represents a truncated T287D
CC   mutant Aurora A kinase sequence which was prepared and crystallised in an
CC   example from the invention. This sequence comprises the Aurora A kinase
CC   domain and also contains a vector-encoded sequence at the N-terminus
XX
SQ   Sequence 309 AA;

  Query Match             71.2%;  Score 1503;  DB 6;  Length 309;
  Best Local Similarity   99.7%;  Pred. No. 1.4e-130;
  Matches  287;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        114 LASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAG 173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 LASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAG 81

Qy        174 VEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRT 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 VEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRT 141

Qy        234 ATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTL 293
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        142 ATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTDLCGTL 201

Qy        294 DYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVT 353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 DYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVT 261

Qy        354 EGARDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASK 401
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 EGARDLISRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASK 309


RESULT 23
ABP97473
ID   ABP97473 standard; protein; 300 AA.
XX
AC   ABP97473;
XX
DT   19-SEP-2003  (first entry)
XX
DE   Truncated mutant human Aurora A kinase, [T287] Aurora A (122-400).
XX
KW   Aurora A kinase; human; catalytic domain; chromosome 20q13; 3D structure;
KW   X-ray crystallography; protein co-ordinate data; serine/threonine kinase;
KW   cell cycle regulation; G2/M boundary; oncogenesis;
KW   hyperproliferative disorder; cancer; tumour; breast; colon; colorectal;
KW   Aurora homologue design; rational drug design; drug screening; enzyme;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .21
FT                   /note= "pET28a vector-encoded sequence, comprising a
FT                   hexahistidine sequence"
FT   Domain          24. .108
FT                   /label= N-terminal_domain
FT                   /note= "Corresponds to residues 125-208 of the native
FT                   human Aurora A kinase"
FT   Active-site     109. .114
FT                   /label= ATP_binding_pocket
FT                   /note= "Corresponds to residues 209-214 of the native
FT                   human Aurora A kinase. Also referred to as a hinge region
FT                   between the N- and C-terminal domains"
FT   Binding-site    110
FT                   /note= "The peptide carbonyl group of this residue binds
FT                   adenine nitrogen atom N6. Corresponds to residue 210 of
FT                   the native human Aurora A kinase"
FT   Binding-site    112
FT                   /note= "The main chain nitrogen of this residue binds
FT                   adenine nitrogen atom N1. Corresponds to residue 212 of
FT                   the native human Aurora A kinase"
FT   Domain          115. .274
FT                   /label= C-terminal_domain
FT                   /note= "Corresponds to residues 215-374 of the native
FT                   human Aurora A kinase"
FT   Misc-difference 187
FT                   /note= "Asp replaces wild-type Thr. Corresponds to
FT                   position 287 of the native human Aurora A kinase"
XX
PN   WO2003031606-A2.
XX
PD   17-APR-2003.
XX
PF   08-OCT-2002; 2002WO-GB004589.
XX
PR   10-OCT-2001; 2001GB-00024299.
XX
PA   (ASTR ) ASTRAZENECA AB.
PA   (ASTR ) ASTRAZENECA UK LTD.
XX
PI   Anderson M,  Keen NJ,  Pannifer ADB,  Pauptit RA,  Rowsell S;
XX
DR   WPI; 2003-403104/38.
XX
PT   Novel crystalline form of polypeptide comprising catalytic domain of
PT   Aurora kinase, whose atomic coordinate data is useful for designing or
PT   selecting Aurora chemical inhibitor, and for designing an Aurora protein.
XX
PS   Example 1; Page 119; 137pp; English.
XX
CC   The invention relates to the three-dimensional structure of human Aurora
CC   A kinase. The invention describes the methods used to prepare and
CC   crystallise truncated mutant forms of human Aurora A kinase comprising
CC   the catalytic domain, which were then subjected to X-ray crystallography
CC   to generate atomic coordinates. The mutant forms of Aurora A used in the
CC   invention all contain a T287D substitution, which produces a mimic of the
CC   activated protein which can be provided as a homogeneous sample (native
CC   Aurora A kinase is activated by phosphorylation of T287). Two crystalline
CC   forms of the Aurora A kinase catalytic domain were produced - one
CC   complexed with the hydrolysable ATP analogue ANP-PNP, and one complexed
CC   with a synthetic inhibitor. The ATP analogue was found to occupy a cleft
CC   between the N-terminal domain (residues 125-208) and the C-terminal
CC   domain (residues 215-374). Aurora A kinase is a serine/threonine kinase
CC   which is involved in cell cycle regulation, with its expression and
CC   activity peaking at the G2/M boundary. It has been implicated in
CC   oncogenesis - the Aurora A gene maps to chromosome 20q13, a region which
```

```
CC   is frequently amplified in human tumours including breast and colon
CC   tumours, and it is overexpressed in over 50% of primary colorectal
CC   tumours. The crystal structure of human Aurora A kinase may be used to
CC   select or design chemical modulators of Aurora kinase, particularly
CC   Aurora kinase inhibitors. These modulators may be used to prevent or
CC   treat undesirable physical and pharmacological consequences of
CC   inappropriate Aurora activity, especially cancers or other
CC   hyperproliferative disorders. The Aurora A kinase 3D structure may also
CC   be used in designing an Aurora protein or Aurora homologue, and in
CC   elucidating the 3D structure of other proteins with structural similarity
CC   to Aurora kinases. The present sequence represents a truncated T287D
CC   mutant Aurora A kinase sequence which was prepared and crystallised in an
CC   example from the invention. This sequence comprises the Aurora A kinase
CC   domain and also contains a vector-encoded sequence at the N-terminus
XX
SQ   Sequence 300 AA;

  Query Match            69.2%;  Score 1460;  DB 6;  Length 300;
  Best Local Similarity  99.6%;  Pred. No. 1.3e-126;
  Matches  278;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        123 SKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREV 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 SKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREV 81

Qy        183 EIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELAN 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 EIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELAN 141

Qy        243 ALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIE 302
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        142 ALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTDLCGTLDYLPPEMIE 201

Qy        303 GRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISR 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISR 261

Qy        363 LLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASK 401
              |||||||||||||||||||||||||||||||||||||||
Db        262 LLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASK 300


RESULT 24
ADK71858
ID   ADK71858 standard; protein; 248 AA.
XX
AC   ADK71858;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human kinase and phosphatase KPP-35 protein.
XX
KW   human; kinase; phosphatase; KPP; cardiovascular; antiarteriosclerotic;
KW   hypotensive; vasotropic; antiinflammatory; antianginal; anti-HIV;
KW   antiallergic; antiasthmatic; immunosuppressive; antithyroid;
KW   dermatological; antidiabetic; nephrotropic; antigout; gastrointestinal;
KW   neuroprotective; osteopathic; antiarthritic; uropathic; ophthalmological;
KW   antirheumatic; antiparkinsonian; nootropic; anticonvulsant; hepatotropic;
KW   antipsoriatic; haemostatic; cytostatic; antilipaemic; antiparasitic;
KW   antihelmintic; antibacterial; virucide; protozoacide; fungicide;
KW   cardiovascular disease; immune system; neurological; growth; development;
KW   cell proliferation; viral; bacterial; fungal; parasitic; protozoan;
KW   helminthic infection; transgenic; gene therapy; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO2004018641-A2.
XX
PD   04-MAR-2004.
XX
PF   25-AUG-2003; 2003WO-US026635.
XX
PR   26-AUG-2002; 2002US-0406172P.
PR   25-SEP-2002; 2002US-0413910P.
PR   27-SEP-2002; 2002US-0414296P.
PR   11-OCT-2002; 2002US-0417821P.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Baughn MR,  Richardson TW,  Marquis JP,  Swarnakar A,  Tang YT;
PI   Becha SD,  Emerling BM,  Jin P,  Wilson AD,  Yue H,  Gietzen KJ;
PI   Chang H,  Yang YG,  Lee SY,  Khare R,  Elliott VS,  Hafalia AJA;
PI   Chawla NK,  Ramkumar J,  Gururajan R,  Tribouley CM,  Chien D,  Tran UK;
PI   Murage J;
XX
DR   WPI; 2004-226830/21.
DR   N-PSDB; ADK71917.
XX
PT   New human kinases and phosphatases, useful for diagnosing, treating or
PT   preventing atherosclerosis, hypertension, AIDS, allergy, multiple
PT   sclerosis, osteoarthritis, Alzheimer's disease, Crohn's disease, cancer
PT   or hepatitis.
XX
PS   Claim 1; SEQ ID NO 35; 347pp; English.
XX
CC   The invention relates to a novel isolated polypeptide which is a human
CC   kinase and phosphatase (KPP). The polypeptide of the invention
CC   demonstrates cardiovascular, antiarteriosclerotic, hypotensive,
CC   vasotropic, antiinflammatory, antianginal, anti-HIV, antiallergic,
CC   antiasthmatic, immunosuppressive, antithyroid, dermatological,
CC   antidiabetic, nephrotropic, antigout, gastrointestinal, neuroprotective,
CC   osteopathic, antiarthritic, uropathic, ophthalmological, antirheumatic,
CC   antiparkinsonian, nootropic, anticonvulsant, hepatotropic, antipsoriatic,
CC   haemostatic, cytostatic, antilipaemic, antiparasitic, antihelmintic,
CC   antibacterial, virucide, protozoacide and fungicide activities. The
CC   kinase and phosphatase (KPP) polynucleotides, polypeptides, agonists and
CC   antagonists may be useful for diagnosing, treating or preventing
CC   disorders such as cardiovascular diseases, immune system disorders,
CC   neurological disorders, disorders affecting growth and development, cell
CC   proliferative disorders and viral, bacterial, fungal, parasitic,
CC   protozoan or helminthic infections. Furthermore, the molecules of the
CC   invention may be useful for creating transgenic animals to model human
CC   disease and during gene therapy. The current sequence is that of a human
CC   KPP protein of the invention.
XX
SQ   Sequence 248 AA;

  Query Match            57.6%;  Score 1215;  DB 8;  Length 248;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-104;
  Matches  235;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTAT 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTAT 235


RESULT 25
AAW18083
ID   AAW18083 standard; protein; 344 AA.
XX
AC   AAW18083;
XX
DT   07-SEP-1997  (first entry)
XX
```

```
DE   Human Aurora-1.
XX
KW   Aurora-1; AUR-1; signal transduction; protein kinase; tumour; cancer;
KW   protein kinase; gene therapy; diagnosis; antibody.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .74
FT                   /label= N-terminal_domain
FT   Domain          74. .344
FT                   /label= Kinase_domain
FT   Modified-site   232
FT                   /label= Phosphorylation
FT                   /note= "cAMP-dependent protein kinase phosphorylation
FT                   site conserved in AUR-2 and yeast and Drosophila
FT                   homologues"
FT   Misc-difference 265
FT                   /note= "deduced residue from some cDNA clones is Gly"
XX
PN   WO9722702-A1.
XX
PD   26-JUN-1997.
XX
PF   25-NOV-1996;   96WO-US018859.
XX
PR   18-DEC-1995;   95US-0008809P.
PR   14-AUG-1996;   96US-0023943P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Mossie KG;
XX
DR   WPI; 1997-341693/31.
DR   N-PSDB; AAT67289.
XX
PT   Aurora-1 and Aurora-2 and related genes - useful in tumour gene therapy.
XX
PS   Claim 3; Page 82-83; 98pp; English.
XX
CC   Novel human proteins, termed Aurora-1 (AAW18083) and Aurora-2 (AAW18084)
CC   (AUR-1 and AUR-2), are related serine/threonine kinases with short N-
CC   terminal extensions that appear to be involved in cancer and/or signal
CC   transduction disorders. Their amino acid sequences were deduced from
CC   pancreatic tumour cDNA clones (AAT67289- 90). AUR-1 and AUR-2 appear to
CC   regulate nuclear division, with disruption of their signaling resulting
CC   in polyploid cells. AUR-1 RNA is broadly expressed in rapidly dividing
CC   cells from normal and tumour tissues. AUR polypeptides can be expressed
CC   in host cells and used to raise diagnostic antibodies and to screen for
CC   compounds that interact with AUR-1 and/or AUR-2
XX
SQ   Sequence 344 AA;

  Query Match             50.1%;  Score 1057.5;  DB 2;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:||||||| | :|:|||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 26
AAY22475
ID   AAY22475 standard; protein; 344 AA.
XX
AC   AAY22475;
XX
DT   29-SEP-1999  (first entry)
XX
DE   Human AUR1 protein sequence.
XX
KW   AUR1; AUR2; human; AUR modulator; cancer; glioma; medullablastoma;
KW   chondrosarcoma; pancreatic tumour; proliferative disease; diagnosis;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9937788-A2.
XX
PD   29-JUL-1999.
XX
PF   21-JAN-1999;   99WO-US001283.
XX
PR   22-JAN-1998;   98US-00012135.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Mossie K;
XX
DR   WPI; 1999-458699/38.
DR   N-PSDB; AAX99724.
XX
PT   New nucleic acid encoding human AUR1 and 2 polypeptides, used to identify
PT   specific modulators for treating cancer or for diagnosis.
XX
PS   Claim 11; Page 140-141; 153pp; English.
XX
CC   This sequence is the human AUR1 protein of the invention. The AUR1 and
CC   AUR2 proteins can be used to identify specific modulators of, and to
CC   generate specific antibodies recognising AUR1 and AUR2. The modulators
CC   can be used for treating conditions involving abnormal AUR signal
CC   transduction, specifically cancer (of colon, breast, kidney, ovary,
CC   bladder, head or neck, also glioma, medullablastoma, chondrosarcoma and
CC   pancreatic tumours, particularly of colon (specifically), breast or
CC   kidney). The modulators can also be used for studying their effects in
CC   animal models of proliferative disease. Probes, based on the coding
CC   sequences are used, diagnostically, to detect or quantify AUR mRNA, by
CC   hybridisation or polymerase chain reaction (PCR). The DNA, optionally
CC   mutated, are useful in gene therapy. Ab are used as diagnostic
CC   immunoassay reagents for detecting the proteins
XX
SQ   Sequence 344 AA;

  Query Match             50.1%;  Score 1057.5;  DB 2;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
```

```
Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:|||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 27
AAG67615
ID   AAG67615 standard; protein; 344 AA.
XX
AC   AAG67615;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human protein.
XX
KW   Human; protein kinase; protein phosphatase; signal transduction.
XX
OS   Homo sapiens.
XX
PN   WO200109316-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-JP005061.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   18-OCT-1999;   99US-0159590P.
PR   11-JAN-2000; 2000JP-00118776.
PR   17-FEB-2000; 2000US-0183322P.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T,  Funahashi S;
PI   Senoo C,  Nezu J;
XX
DR   WPI; 2001-570286/64.
XX
PT   New genes encoding proteins with protein kinase/protein phosphatase
PT   activity, useful in the diagnosis and treatment of diseases.
XX
PS   Example 4; Page 97-100; 233pp; Japanese.
XX
CC   The specification describes human protein kinase/protein phosphatases. It
CC   is expected that the protein kinase/protein phosphatase gene participates
CC   in signal transduction in cells. The protein kinase/protein phosphatase
CC   polypeptides and polynucleotides are useful for developing diagnostics
CC   and treatment agents for human and animal diseases. The protein
CC   kinase/protein phosphatase polypeptides are useful as target molecules in
CC   designing novel drugs. The protein kinase/protein phosphatase
CC   polynucleotides are useful as a source of probes and primers, which may
CC   be used to isolate homologous sequences. The present sequence represents
CC   a human protein, which is used in the course of the invention
XX
SQ   Sequence 344 AA;

  Query Match           50.1%;  Score 1057.5;  DB 4;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:|||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 28
AAG67436
ID   AAG67436 standard; protein; 344 AA.
XX
AC   AAG67436;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human polypeptide.
XX
KW   Human; protein kinase; protein phosphatase; signal transduction;
KW   intracellular signalling pathway.
XX
OS   Homo sapiens.
XX
PN   WO200109345-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-JP005060.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   18-OCT-1999;   99US-0159590P.
PR   11-JAN-2000; 2000JP-00118776.
PR   17-FEB-2000; 2000US-0183322P.
PR   02-MAY-2000; 2000JP-00183767.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T,  Funahashi S;
PI   Senoo C,  Nezu J;
XX
DR   WPI; 2001-564736/63.
XX
PT   New genes encoding protein kinase and protein phosphatase, useful for
PT   identifying modulators which can be used to treat human or animal
PT   disorders associated with the expression or function of these enzymes.
XX
PS   Example 4; Page 199-201; 336pp; Japanese.
XX
CC   The specification describes human protein kinase/protein phosphatases.
CC   The polypeptides are expected to participate in signal transduction in
CC   cells. The kinase phosphatases are connected with intracellular
CC   signalling pathways. Antisense oligonucleotides and compounds identified
CC   by screening (agonists or antagonists) can be used to treat human or
CC   animal disorders associated with the expression or function of the
```

```
CC   protein. In addition, the polypeptides may be used as target molecules
CC   fir drug development. The present sequence represents a polypeptide, used
CC   in the course of the invention
XX
SQ   Sequence 344 AA;

  Query Match             50.1%;  Score 1057.5;  DB 4;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||   |  ||:||||:||:|||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 29
ABR48188
ID   ABR48188 standard; protein; 344 AA.
XX
AC   ABR48188;
XX
DT   12-JUN-2003  (first entry)
XX
DE   Human bladder cancer associated protein sequence SEQ ID NO:93.
XX
KW   Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003003906-A2.
XX
PD   16-JAN-2003.
XX
PF   03-JUL-2002; 2002WO-US021338.
XX
PR   03-JUL-2001; 2001US-0302814P.
PR   03-AUG-2001; 2001US-0310099P.
PR   08-NOV-2001; 2001US-0343705P.
PR   13-NOV-2001; 2001US-0350666P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Mack DH,  Aziz N;
XX
DR   WPI; 2003-201532/19.
DR   N-PSDB; ACC51000.
XX
PT   Detecting a bladder cancer-associated transcript in a cell from a
PT   patient, comprises contacting a biological sample from the patient with a
PT   bladder cancer-associated polynucleotide or antibody.
XX
PS   Claim 10; Page 265; 307pp; English.
XX
CC   The present invention describes a method for detecting a bladder cancer-
CC   associated transcript in a cell from a patient. The method comprises
CC   contacting a biological sample from the patient with a polynucleotide
CC   that selectively hybridises to a sequence that is 80 % identical to a
CC   table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC   encode the human bladder cancer-associated proteins given in ABR48146 to
CC   ABR48242). Bladder cancer-associated sequences from the present invention
CC   have cytostatic activities, and can be used in antisense gene therapy and
CC   in vaccine production. The method can be used for detecting a bladder
CC   cancer-associated transcript in a cell from a patient. The method is
CC   useful in diagnosing or treating bladder cancer and in screening for
CC   compounds that modulate bladder cancer, such as hormones or antibodies.
CC   The nucleic acid molecules from the present invention may be used in
CC   various screening and diagnostic methods, and for gene therapy, vaccine
CC   and/or antisense/inhibition applications
XX
SQ   Sequence 344 AA;

  Query Match             50.1%;  Score 1057.5;  DB 6;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||   |  ||:||||:||:|||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 30
ABU56641
ID   ABU56641 standard; protein; 344 AA.
XX
AC   ABU56641;
XX
DT   02-APR-2003  (first entry)
XX
DE   Lung cancer-associated polypeptide #234.
XX
KW   Lung cancer-associated polypeptide; cytostatic; emphysema;
KW   antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW   small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW   chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW   interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS   Unidentified.
XX
PN   WO200286443-A2.
XX
PD   31-OCT-2002.
XX
PF   18-APR-2002; 2002WO-US012476.
XX
PR   18-APR-2001; 2001US-0284770P.
```

PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
DR N-PSDB; ABX76370.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 27; Page 371; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC invention
XX
SQ Sequence 344 AA;

Query Match 50.1%; Score 1057.5; DB 6; Length 344;
Best Local Similarity 63.2%; Pred. No. 3.7e-89;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||:| ||:|||||||:|:|| ||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
```

RESULT 31
ABR92152

ID ABR92152 standard; protein; 344 AA.
XX
AC ABR92152;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human cervical cancer cell marker protein SEQ ID NO:214.
XX
KW Human; cervical cancer; cervical cancer marker; cancer therapy;
KW detection; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2002101075-A2.
XX
PD 19-DEC-2002.
XX
PF 12-JUN-2002; 2002WO-US018638.
XX
PR 13-JUN-2001; 2001US-0298155P.
PR 13-JUN-2001; 2001US-0298159P.
PR 14-NOV-2001; 2001US-0335936P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
PI Gannavarapu M, Glatt K, Hoersch S;
XX
DR WPI; 2003-156967/15.
DR N-PSDB; ACF12935.
XX
PT New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
PS Claim 4; Page 366-367; 386pp; English.
XX
CC ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (M1) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterising,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials
XX
SQ Sequence 344 AA;

Query Match 50.1%; Score 1057.5; DB 6; Length 344;
Best Local Similarity 63.2%; Pred. No. 3.7e-89;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||:| ||:|||||||:|:|| ||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:|||| |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
```

```
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:|||| 
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 32
ADB80537
ID   ADB80537 standard; protein; 344 AA.
XX
AC   ADB80537;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Ovarian cancer-associated protein #51.
XX
KW   cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW   post-operative chemotherapy; radiation therapy; tumour prognosis;
KW   pre-cancerous lesion detection.
XX
OS   Homo sapiens.
XX
PN   WO2002102235-A2.
XX
PD   27-DEC-2002.
XX
PF   18-JUN-2002; 2002WO-US019297.
XX
PR   18-JUN-2001; 2001US-0299234P.
PR   27-AUG-2001; 2001US-0315287P.
PR   05-SEP-2001; 2001US-0317544P.
PR   13-NOV-2001; 2001US-0350666P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Mack DH,  Gish KC;
XX
DR   WPI; 2003-167431/16.
DR   N-PSDB; ADB80536.
XX
PT   Detecting an ovarian cancer-associated transcript in a cell from a
PT   patient, comprises contacting a biological sample from the patient with a
PT   polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS   Claim 13; Page 310; 332pp; English.
XX
CC   The invention relates to a method of detecting an ovarian cancer-
CC   associated transcript in a cell from a patient, by contacting a
CC   biological sample from the patient with a polynucleotide that selectively
CC   hybridizes to a sequence at least 80% identical to any of one of 80
CC   nucleic acid sequences given in the specification. The method is useful
CC   in diagnosing ovarian cancer and in identifying and using agents and/or
CC   targets that inhibit ovarian cancer. The nucleic acid molecule,
CC   polypeptide and the antibody may also be used in detecting ovarian
CC   cancers, monitoring and early detection of relapse following treatment,
CC   monitoring response to therapy, selecting patients for post-operative
CC   chemotherapy or radiation therapy, in selecting mode of therapy,
CC   determining tumour prognosis, early detection of pre-cancerous lesions,
CC   and as vaccines. This sequence corresponds to one of the proteins used
CC   for the detection method of the invention.
XX
SQ   Sequence 344 AA;

  Query Match           50.1%;  Score 1057.5;  DB 7;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:|||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:|||| 
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 33
ADN38885
ID   ADN38885 standard; protein; 344 AA.
XX
AC   ADN38885;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:203.
XX
KW   Human; differential expression; cancer; angiogenic disorder;
KW   fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW   inflammatory disease; autoimmune disease;
KW   retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW   detection; diagnosis; prognosis; drug screening; drug targeting;
KW   wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW   vulnerary; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003042661-A2.
XX
PD   22-MAY-2003.
XX
PF   13-NOV-2002; 2002WO-US036810.
XX
PR   13-NOV-2001; 2001US-0350666P.
PR   21-NOV-2001; 2001US-0332464P.
PR   29-NOV-2001; 2001US-0334393P.
PR   03-DEC-2001; 2001US-0335394P.
PR   14-DEC-2001; 2001US-0340376P.
PR   08-JAN-2002; 2002US-0347211P.
PR   10-JAN-2002; 2002US-0347349P.
PR   08-FEB-2002; 2002US-0355250P.
PR   13-FEB-2002; 2002US-0356714P.
PR   20-FEB-2002; 2002US-0359077P.
PR   29-MAR-2002; 2002US-0368809P.
PR   04-APR-2002; 2002US-0370110P.
PR   12-APR-2002; 2002US-0372246P.
PR   05-JUN-2002; 2002US-0386614P.
PR   16-JUL-2002; 2002US-0396839P.
PR   22-JUL-2002; 2002US-0397775P.
PR   22-JUL-2002; 2002US-0397845P.
PR   09-SEP-2002; 2002US-0409450P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D,  Aziz N,  Ginsburg WM,  Gish KC,  Glynne R,  Hevezi PA;
```

```
PI   Mack DH,  Murray R,  Watson SR,  Wilson KE,  Zlotnik A;
XX
DR   WPI; 2003-468649/44.
DR   N-PSDB; ADN38884.
XX
PT   Determining the presence or absence of a pathological cell in a patient,
PT   useful for diagnosing, prognosing or treating cancer, comprises detecting
PT   a nucleic acid in a biological sample.
XX
PS   Claim 12; SEQ ID NO 203; 1385pp; English.
XX
CC   The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC   whose expression is upregulated or downregulated in specific cancers or
CC   other diseases such as angiogenic or fibrotic disorders, and to methods
CC   of determining the presence or absence of a pathological cell in a
CC   patient by detecting a nucleic acid at least 80% identical to those of
CC   the invention or by detecting a polypeptide of the invention. The
CC   invention also relates to expression vectors and host cells comprising a
CC   nucleic acid of the invention; antibodies which specifically bind a
CC   polypeptide of the invention; use of such antibodies for drug targeting;
CC   and methods of screening for modulators of activity or expression of the
CC   polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC   antibodies and methods are useful for diagnosing, prognosing and treating
CC   cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC   atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC   neovascularistaion syndromes, scarring and uterine fibroids. They may
CC   also be useful in wound healing and in contraception. The present
CC   sequence represents a polypeptide of the invention.
XX
SQ   Sequence 344 AA;

  Query Match             50.1%;  Score 1057.5;  DB 7;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:||||||| | :|:|||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 34
ADN95464
ID   ADN95464 standard; protein; 344 AA.
XX
AC   ADN95464;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human BEC/LEC-related protein sequence SeqID387.
XX
KW   growth; differentiation; blood endothelial cell; BEC;
KW   lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW   lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
KW   vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;
KW   inflammatory disease; cancer metastasis; lymphatic system; human.
XX
OS   Homo sapiens.
XX
PN   WO2003080640-A1.
XX
PD   02-OCT-2003.
XX
PF   07-MAR-2003; 2003WO-US006900.
XX
PR   07-MAR-2002; 2002US-0363019P.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (LICN ) LICENTIA LTD.
XX
PI   Alitalo K,  Makinen T,  Petrova T,  Saharinen P,  Saharinen J;
XX
DR   WPI; 2003-876899/81.
DR   N-PSDB; ADN95465.
XX
PS   Example 1; SEQ ID NO 387; 176pp; English.
XX
CC   This invention relates to a method of differentially modulating the
CC   growth or differentiation of blood endothelial cells (BEC) or lymphatic
CC   endothelial cells (LEC) comprises contacting endothelial cells with a
CC   composition comprising an agent that differentially modulates blood or
CC   lymphatic endothelial cells. Treating hereditary lymphoedema comprises
CC   identifying a human subject with lymphoedema and with a mutation in at
CC   least one allele of a gene encoding a LEC protein, where the mutation
CC   correlates with lymphoedema in human subjects, and with the proviso that
CC   the LEC protein is not VEGFR-3; and administering to the subject a
CC   composition comprising a lymphatic growth agent selected from VEGF-C or
CC   VEGF-D polypeptides and polynucleotides. The invention may be useful for
CC   the development of compounds with an antiangiogenic, cytostatic,
CC   vasotropic or antiinflammatory activity or for gene therapy. The method
CC   is useful in modulating the growth or differentiation of blood
CC   endothelial cells or lymphatic endothelial cells, in treating hereditary
CC   lymphoedema, in screening for an endothelial cell disorder or
CC   predisposition to the disorder or in monitoring the efficacy or toxicity
CC   of a drug on endothelial cells. The agent is useful in manufacturing a
CC   medicament for the differential modulation of blood vessel endothelial
CC   cell or lymphatic vessel endothelial cell growth or differentiation. The
CC   lymphatic growth agent may also be used in manufacturing a medicament for
CC   the treatment of hereditary lymphoedema resulting from a mutation in a
CC   LEC gene or of other diseases involving the lymphatic vessels, such as
CC   various inflammatory diseases and cancer metastasis via the lymphatic
CC   system. The present sequence is that of a human LEC/BEC differentially
CC   expressed protein which is related to the method of the invention.Note:
CC   This sequence does not appear in the specification but was obtained by
CC   the indexer using the source data given in table 14 of the specification.
XX
SQ   Sequence 344 AA;

  Query Match             50.1%;  Score 1057.5;  DB 7;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:||||||| | :|:|||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
```

```
Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 35
ADK67718
ID   ADK67718 standard; protein; 344 AA.
XX
AC   ADK67718;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human modifier of AXIN (MAX) polypeptide.
XX
KW   Human; modifier of AXIN; MAX; cytostatic; gene therapy; protein kinase;
KW   enzyme.
XX
OS   Homo sapiens.
XX
PN   WO2004013308-A2.
XX
PD   12-FEB-2004.
XX
PF   06-AUG-2003; 2003WO-US024560.
XX
PR   06-AUG-2002; 2002US-0401534P.
PR   16-SEP-2002; 2002US-0411153P.
XX
PA   (EXEL-) EXELIXIS INC.
XX
PI   Gendreau SB,  Dora EG,  Lickteig K,  Amundsen CD;
XX
DR   WPI; 2004-157122/15.
DR   N-PSDB; ADK67711.
XX
PT   Identifying a candidate AXIN pathway-modulating agent, useful in
PT   diagnosing and treating cancer, comprises providing an assay system
PT   comprising a MAX polypeptide or nucleic acid.
XX
PS   Example 1; SEQ ID NO 9; 74pp; English.
XX
CC   The present sequence is that of a human modifier of AXIN (MAX)
CC   polypeptide characterised as a serine/threonine kinase. Genetic screens
CC   were designed to identify modifiers of the axin pathway in Caenorhabditis
CC   elegans, where a reduction of function pry-1 (axin) mutant was used.
CC   Genes causing altered phenotypes were isolated, and their human
CC   homologues, termed MAX, were identified, including the present sequence.
CC   These MAX polynucleotides and polypeptides are attractive targets for the
CC   treatment of pathologies associated with a defective AXIN signalling
CC   pathway, such as cancer. Modulation of MAX or its binding partner is
CC   useful for understanding the association of the AXIN pathway and its
CC   members in normal and disease conditions and for developing diagnostic
CC   and therapeutic modalities for AXIN related pathologies. MAX-modulating
CC   agents that act by inhibiting or enhancing MAX expression, directly or
CC   indirectly, e.g. by affecting MAX function can be identified using
CC   methods of the invention. MAX modulating agents are useful in diagnosis,
CC   therapy and pharmaceutical development. Preferred MAX modulating agents
CC   include antisense and phosphorothioate morpholino oligomers.
XX
SQ   Sequence 344 AA;

  Query Match             50.1%;  Score 1057.5;  DB 8;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |       |:|: |||      :| :   |  ||  |:         | :

Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||| || |:|||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 36
ADQ89834
ID   ADQ89834 standard; protein; 344 AA.
XX
AC   ADQ89834;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Antagonist of cell cycle progression polypeptide #132.
XX
KW   Cytostatic; cancer; cell division cycle; mitosis; meiosis;
KW   cell cycle progression.
XX
OS   Homo sapiens.
XX
PN   WO2004063362-A2.
XX
PD   29-JUL-2004.
XX
PF   31-DEC-2003; 2003WO-GB005635.
XX
PR   10-JAN-2003; 2003US-0439123P.
PR   06-MAY-2003; 2003US-0468402P.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Glover D,  Bell G,  Frenz L,  Midgley C;
XX
DR   WPI; 2004-544089/52.
DR   N-PSDB; ADQ89833.
XX
PT   New cell cycle progression genes and proteins for modulating cell cycle
PT   progression in cells, for preventing, treating or diagnosing cell
PT   proliferative diseases (e.g. cancer) or for identifying modulators of
PT   mitosis or meiosis.
XX
PS   Claim 2; SEQ ID NO 264; 461pp; English.
XX
CC   The present invention relates to a polynucleotide for preventing,
CC   treating or diagnosing a disease in an individual. The composition or the
CC   polypeptide, polynucleotide or RNA precursor, or antibody is useful for
CC   diagnosing, preventing or treating diseases (e.g. cell proliferative
CC   diseases such as cancer) in an individual. These may also be used for
CC   identifying substances capable of binding to or modulating the function
CC   of the polypeptide, capable of affecting the function of the
CC   corresponding gene, or capable of inhibiting the cell division cycle or
CC   cell cycle progression, preferably mitosis and/or meiosis. The present
CC   sequence represents an antagonist of cell cycle progression protein
CC   sequence.
XX
```

```
SQ   Sequence 344 AA;

  Query Match           50.1%;  Score 1057.5;  DB 8;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::||||||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 37
ABM81826
ID   ABM81826 standard; protein; 344 AA.
XX
AC   ABM81826;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Tumour-associated antigenic target (TAT) polypeptide PRO70812, SEQ:4700.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD,  Zhang Z,  Zhou Y;
XX
DR   WPI; 2004-347921/32.
DR   N-PSDB; ACN40117.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 12; SEQ ID NO 4700; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT polypeptide of the invention
XX
SQ   Sequence 344 AA;

  Query Match           50.1%;  Score 1057.5;  DB 8;  Length 344;
  Best Local Similarity   63.2%;  Pred. No. 3.7e-89;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::||||||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
              :|  | :|  |||: ||| :   ||  |:
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342


RESULT 38
ABR61580
ID   ABR61580 standard; protein; 344 AA.
XX
AC   ABR61580;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human HsAIRK-2 (Aurora A) protein.
XX
KW   RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; HsAIRK-2;
KW   Aurora A.
XX
OS   Homo sapiens.
XX
PN   WO2003087395-A2.
XX
PD   23-OCT-2003.
XX
PF   15-APR-2003; 2003WO-IB002972.
XX
```

PR 15-APR-2002; 2002US-0372483P.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Garbay C, Gigoux V, Camonis J, L'hoste S, Samson J;
XX
DR WPI; 2003-845340/78.
DR N-PSDB; ACF58069.
XX
PT Identifying anti-tumoral compounds, comprises determining the capacity of
PT a compound to inhibit interaction between RasGAP and Drosophila
PT melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and
PT aurora binding protein.
XX
PS Disclosure; Page 62-63; 0pp; English.
XX
CC The invention relates to identifying a biologically active compound with
CC anti-tumoural properties, where the compound is studied for its capacity
CC to inhibit the interaction between: (a) RasGAP and the Drosophila
CC melanogaster Aurora kinase or an orthologue of the kinase, or its
CC fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding
CC protein and the D. melanogaster Aurora kinase or an orthologue of the
CC kinase, or its fragment. The method is useful for identifying anti-
CC tumoural agents. The present sequence represents a human HsAIRK-2 (Aurora
CC A) protein, an orthologue of the D. melanogaster Aurora kinase
XX
SQ Sequence 344 AA;

Query Match 50.1%; Score 1056.5; DB 7; Length 344;
Best Local Similarity 63.2%; Pred. No. 4.6e-89;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193

Qy        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309
Db        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253

Qy        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369
Db        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPMGAQDLISKLLRHNPS 313

Qy        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
Db        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
```

RESULT 39
AAW99783
ID AAW99783 standard; protein; 343 AA.
XX
AC AAW99783;
XX
DT 07-JUN-1999 (first entry)
XX
DE Rat AIM-1.
XX
KW Rat; AIM-1; aurora and IPL-1 like midbody-associated protein kinase;
KW cell cycle regulating protein; serine-threonine kinase;
KW cell proliferation; cancer.
XX
OS Rattus sp.
XX

PN WO9909160-A1.
XX
PD 25-FEB-1999.
XX
PF 17-AUG-1998; 98WO-JP003641.
XX
PR 15-AUG-1997; 97JP-00235371.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Tatsuka M, Terada Y;
XX
DR WPI; 1999-181033/15.
DR N-PSDB; AAX19715.
XX
PT DNA encoding cell cycle-regulating protein, aurora and IPL-1 like midbody
PT -associated protein kinase AIM-1 - useful for treatment of diseases
PT associated with abnormal proliferation of cells e.g. cancers.
XX
PS Claim 1; Page 31-35; 44pp; Japanese.
XX
CC The present sequence represents rat aurora and IPL-1 like midbody-
CC associated protein kinase, designated AIM-1. AIM-1 protein inhibitors
CC obtained can be used as remedies for treatment of diseases associated
CC with abnormal cell proliferation, e.g. cancers. The genes and AIM-1
CC proteins can be applied in the screening of substances with serine-
CC threonine kinase inhibitory activity. Hybridised oligonucleotides and
CC peptide nucleic acids can be used to the inhibit expression of the AIM-1
CC proteins
XX
SQ Sequence 343 AA;

Query Match 49.8%; Score 1050; DB 2; Length 343;
Best Local Similarity 57.6%; Pred. No. 1.8e-88;
Matches 204; Conservative 51; Mismatches 73; Indels 26; Gaps 4;

```
Qy         36 PLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNT 95
Db          9 PWPYGSKTSQSGL--NTLPQRV-LRKEPAVTPAQALMNR-----------------SNS 47

Qy         96 QKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQS 155
Db         48 QSTAVPGQKLTENKGATALQGSQ-----SRQPFTIDNFEIGRPLGKGKFGNVYLAREKKS 102

Qy        156 KFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPL 215
Db        103 RFIVALKILFKSQIEKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPR 162

Qy        216 GTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADF 275
Db        163 GELYKELQKSGTFDEQRTATIMEELSDALMYCHKKKVIHRDIKPENLLLGLQGELKIADF 222

Qy        276 GWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQ 335
Db        223 GWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHS 282

Qy        336 ETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
Db        283 ETYRRIVKVDLKFPSSMPLGAKDLISKLLKHNPSQRLPLEQVSAHPWVRANSRR 336
```

RESULT 40
AAY27052
ID AAY27052 standard; protein; 347 AA.
XX
AC AAY27052;
XX
DT 08-OCT-1999 (first entry)
XX
DE Human protein kinase (HPKM)-1 (clone ID 2940).
XX
KW Human protein kinase molecule; HPKM; human; protein kinase;
KW phosphate group; cancer; immune disorder.

```
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   60
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   73
FT                   /note= "potential phosphorylation site"
FT   Domain          83. .106
FT                   /note= "potential signature sequence for protein kinase
FT                   catalytic domain"
FT   Modified-site   111
FT                   /note= "potential phosphorylation site"
FT   Modified-site   171
FT                   /note= "potential phosphorylation site"
FT   Modified-site   179
FT                   /note= "potential phosphorylation site"
FT   Domain          199. .211
FT                   /note= "potential signature sequence for protein kinase
FT                   catalytic domain"
FT   Modified-site   230
FT                   /note= "potential phosphorylation site"
FT   Modified-site   235
FT                   /note= "potential phosphorylation site"
FT   Modified-site   280
FT                   /note= "potential phosphorylation site"
FT   Modified-site   282
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   284
FT                   /note= "potential phosphorylation site"
FT   Modified-site   316
FT                   /note= "potential phosphorylation site"
FT   Modified-site   334
FT                   /note= "potential phosphorylation site"
XX
PN   WO9938981-A2.
XX
PD   05-AUG-1999.
XX
PF   12-JAN-1999;   99WO-US000661.
XX
PR   30-JAN-1998;   98US-00016000.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Bandman O,  Hillman JL,  Lal P,  Akerblom IE,  Shah P,  Corley NC;
PI   Guegler KJ;
XX
DR   WPI; 1999-479190/40.
DR   N-PSDB; AAX89850.
XX
PT   New human protein kinase molecules useful for treating or preventing
PT   cancer or an immune disorder.
XX
PS   Claim 1; Page 64-65; 77pp; English.
XX
CC   The invention provides human protein kinase molecules (HPKM) (AAY27052-
CC   57) and nucleic acid sequences (AAX89850-55) encoding the HPKM
CC   polypeptides respectively. The HPKM polypeptides can be produced
CC   recombinantly by standard recombinant methodology. Protein kinases add
CC   phosphate groups to proteins. HPKM polypeptide or an HPKM antagonist is
CC   used to treat or prevent cancer or an immune disorder. The present
CC   sequence represents the amino acid sequence of HPKM-1
XX
SQ   Sequence 347 AA;

  Query Match             49.6%;  Score 1046;  DB 2;  Length 347;
  Best Local Similarity   62.7%;  Pred. No. 4.4e-88;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   26;  Gaps    6;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTAT---YITELANALSY 246
              |||||||  ||:|  |:||||||| | :|:|||   |||||||    : |||:|| |
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATVRAIMEELADALMY 193

Qy        247 CHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMH 306
              || |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||
Db        194 CHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMH 253

Qy        307 DEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKH 366
              :|||||| :|||||| ||| ||||: :: |||:|| :|:  ||   |  ||:||||:||:|
Db        254 NEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRH 313

Qy        367 NPSQRPMLREVLEHPWITANSSK---PSNCQN 395
              |||:|  | :|  |||: ||| :   ||  |:
Db        314 NPSERLPLAQVSAHPWVRANSRRVLPPSALQS 345


RESULT 41
ABU57643
ID   ABU57643 standard; protein; 347 AA.
XX
AC   ABU57643;
XX
DT   09-APR-2003  (first entry)
XX
DE   Differentially expressed breast cancer associated protein #30.
XX
KW   Breast cancer; differential gene expression; BC-cDNA;
KW   breast cancer diagnosis; breast cancer monitoring;
KW   breast cancer treatment; breast cancer staging.
XX
OS   Homo sapiens.
XX
PN   US2002156263-A1.
XX
PD   24-OCT-2002.
XX
PF   04-OCT-2001; 2001US-00974298.
XX
PR   05-OCT-2000; 2000US-0238331P.
XX
PA   (CHEN/) CHEN H.
XX
PI   Chen H;
XX
DR   WPI; 2003-182653/18.
XX
PT   New cDNAs, which are differentially expressed in (metastatic) breast
PT   cancer useful for diagnosing or staging, breast cancer, or for monitoring
PT   the treatment of breast cancer in an individual.
XX
PS   Example; SEQ ID NO 136; 30pp; English.
XX
CC   The invention describes a combination of cDNAs (designated BC-cDNAs),
CC   which are differentially expressed in breast cancer. The combination
CC   includes 152 cDNA sequences, or their complements. The protein encoded by
CC   any of these BC-cDNAs is useful for screening several molecules or
CC   compounds to identify at least one ligand that specifically binds the
CC   protein, producing or preparing polyclonal or monoclonal antibodies, or
CC   purifying antibodies from a sample. The antibodies, which specifically
CC   bind the protein differentially expressed in breast cancer is useful for
CC   detecting the expression of a protein in a sample. The BC-cDNAs are also
CC   useful for diagnosing, monitoring the treatment of, or staging, breast
CC   cancer. This is the amino acid sequence of a differentially expressed
CC   breast cancer associated protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from the US patent office at
CC   seqdata.uspto.gov/sequence.html?DocID=20020156263
```

```
XX
SQ   Sequence 347 AA;

  Query Match             49.6%;  Score 1046;  DB 6;  Length 347;
  Best Local Similarity   62.7%;  Pred. No. 4.4e-88;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   26;  Gaps    6;

Qy         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73

Qy        130 LEDPEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db         74 IDDPEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133

Qy        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTAT---YITELANALSY 246
              |||||||| ||:|  |:|||||||| | :|:|||   ||||||||    : |||:|| |
Db        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATVRAIMEELADALMY 193

Qy        247 CHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMH 306
              || |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||||
Db        194 CHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMH 253

Qy        307 DEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKH 366
              :|||||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||||:||:|
Db        254 NEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRH 313

Qy        367 NPSQRPMLREVLEHPWITANSSK---PSNCQN 395
              |||:|  | :|  |||: ||| :   ||  |:
Db        314 NPSERLPLAQVSAHPWVRANSRRVLPPSALQS 345


RESULT 42
ADO57331
ID   ADO57331 standard; protein; 320 AA.
XX
AC   ADO57331;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Kidney development associated protein seqid 98.
XX
KW   nephrotropic; cardiant; antiinfertility; cerebroprotective;
KW   neuroprotective; muscular; cytostatic; osteopathic; gene therapy;
KW   kidney development; kidney disorder; developmental disorder;
KW   circulatory disorder; hearing disorder; heart defect; infertility;
KW   stroke; mental retardation; muscle defect; proliferative disorder;
KW   bone defect; bone disorder; zebrafish.
XX
OS   Danio rerio.
XX
PN   US2004068763-A1.
XX
PD   08-APR-2004.
XX
PF   28-MAR-2003; 2003US-00403571.
XX
PR   29-MAR-2002; 2002US-0368760P.
XX
PA   (HOPK/) HOPKINS N.
PA   (GOLL/) GOLLING G.
PA   (AMST/) AMSTERDAM A.
PA   (SUNZ/) SUN Z.
XX
PI   Hopkins N,  Golling G,  Amsterdam A,  Sun Z;
XX
DR   WPI; 2004-304692/28.
DR   N-PSDB; ADO57330.
XX
PT   New 459 nucleic acids and encoded polypeptides, useful for diagnosing,
PT   treating or preventing a kidney disorder in an organism, or in screening
PT   for compounds that modulate the development of an organism.
XX
PS   Disclosure; SEQ ID NO 98; 347pp; English.
XX
CC   The invention describes an isolated nucleic acid molecule (I) comprising
CC   a sequence having at least 75% sequence identity to the 459 nucleic acid
CC   sequence of 2808 base pairs (SEQ ID NO: 59) given in the specification,
CC   over at least 600 contiguous base pairs, where the nucleic acid functions
CC   in kidney development. (I) is useful for treating or preventing a kidney
CC   disorder in an organism, where the nucleic acid elicits an alteration in
CC   expression of a 459 nucleic acid sequence in the organism and
CC   subsequently treats or prevents a kidney disorder. The nucleic acid may
CC   also be used in diagnosing, preventing and treating a variety of
CC   mammalian diseases and developmental disorders (e.g. circulatory
CC   disorders, hearing disorders, heart defect, infertility, stroke, mental
CC   retardation, muscle defects, proliferative disorders, or bone defects or
CC   disorders) as well as in screening for compounds that modulate the
CC   development of an organism as a whole or of specific tissues or organs
CC   within that organism. This is the amino acid sequence of a kidney
CC   development associated protein.
XX
SQ   Sequence 320 AA;

  Query Match             49.5%;  Score 1045;  DB 8;  Length 320;
  Best Local Similarity   64.1%;  Pred. No. 4.9e-88;
  Matches  200;  Conservative   40;  Mismatches   62;  Indels   10;  Gaps    3;

Qy         94 NTQKSKQPLPSA------PENNPEEELAS--KQKNEESKKRQWALEDFEIGRPLGKGKFG 145
              | :   |  |||       | ||:    |   :   :|  :  :::||:||||||||||
Db          6 NREPRVQQTPSAGVGPLRVEMNPDTHAVSGPGRVPVKSNSKVLSIDDFDIGRPLGKGKFG 65

Qy        146 NVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATR 205
              |||||||:: | ::|||||||:|: | |||||||||:|||||||||||| | |||| ||
Db         66 NVYLARERKLKVVIALKVLFKSQMVKEGVEHQLRREIEIQSHLRHPNILRFYNYFHDDTR 125

Qy        206 VYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLG 265
              |:|||||| | :|:|||:   ||:||||||: |:::|| ||| |:|||||||||||||
Db        126 VFLILEYAPRGEMYKELQRYGHFDDQRTATYMEEVSDALQYCHEKKVIHRDIKPENLLLG 185

Qy        266 SAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVG 325
              |||||||||||||||| || |:|||||||||||||  |||||||||:|||||| |||
Db        186 YRGELKIADFGWSVHAPSLRRRTMCGTLDYLPPEMIEGHSHDEKVDLWSIGVLCYECLVG 245

Qy        326 KPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITA 385
              ||||   : ||||||::|:  ||  |:|||||||:||:|:|| |  || |:||  : |
Db        246 NPPFETRQHAETYKRITKVDLQFPKLVSEGARDLISKLLRHSPSMRLPLRSVMEHRGVKA 305

Qy        386 NSSK--PSNCQN 395
              || :  |  | :
Db        306 NSRRVLPPVCSS 317


Search completed: June 20, 2005, 20:04:24
Job time : 149.005 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:55:54 ; Search time 34.5274 Seconds
(without alignments)
1123.030 Million cell updates/sec

Title: US-10-734-126-4
Perfect score: 2110
Sequence: 1 MDRSKENCISGPVKATAPVG.....TANSSKPSNCQNKESASKQS 403

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2095	99.3	403	2	JC5974	aurora-related kin
2	1654	78.4	395	2	JC5975	aurora-related kin
3	1310.5	62.1	389	2	S52242	protein kinase (EC
4	1295.5	61.4	407	2	S52243	p46Eg265 protein -
5	1049	49.7	345	2	JC4665	protein kinase (EC

ALIGNMENTS

RESULT 1
JC5974
aurora-related kinase 1 (EC 2.7.-.-) - human
C;Species: Homo sapiens (man)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 16-Aug-2004
C;Accession: JC5974
R;Shindo, M.; Nakano, H.; Kuroyanagi, H.; Shirasawa, T.; Mihara, M.; Gilbert, D.J.; Jenk
Biochem. Biophys. Res. Commun. 244, 285-292, 1998
A;Title: cDNA cloning, expression, subcellulard localization, and chromosomal assignment
A;Reference number: JC5974; MUID:98183439; PMID:9514916
A;Accession: JC5974
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-403 <SHI>
A;Cross-references: GB:AF008551
C;Superfamily: protein kinase homology
C;Keywords: phosphotransferase
F;131-383/Domain: protein kinase homology <KIN>

Query Match 99.3%; Score 2095; DB 2; Length 403;
Best Local Similarity 99.3%; Pred. No. 1.2e-78;
Matches 400; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRIPLQ 60

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 AQKLVSSHKPVQNQKQKQLQQTSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLRR 180

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
              |||||||||||||||||||||||||||||||||||||||||||
Db        361 SRLLKHNPSQRPMLREVLEHPWITANSSKPSNCQNKESASKQS 403
```

RESULT 2
JC5975
aurora-related kinase 1 (EC 2.7.-.-) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 16-Aug-2004
C;Accession: JC5975
R;Shindo, M.; Nakano, H.; Kuroyanagi, H.; Shirasawa, T.; Mihara, M.; Gilbert, D.J.; Jenk
Biochem. Biophys. Res. Commun. 244, 285-292, 1998
A;Title: cDNA cloning, expression, subcellulard localization, and chromosomal assignment
A;Reference number: JC5974; MUID:98183439; PMID:9514916
A;Accession: JC5975
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-395 <SHI>
C;Superfamily: protein kinase homology
C;Keywords: phosphotransferase
F;123-374/Domain: protein kinase homology <KIN>

Query Match 78.4%; Score 1654; DB 2; Length 395;
Best Local Similarity 82.7%; Pred. No. 8.7e-61;
Matches 334; Conservative 17; Mismatches 43; Indels 10; Gaps 7;

```
Qy          1 MDRSKENCISGPVKATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQ 60
              ||| |||| :| ||| | |  |||||||| : | ||    :|||||||||||| ||||| |
Db          1 MDRCKENCVSRPVKTTVPF-GPKRVLVTEQIPSQNLGSASSGQAQRVLCPSN-SQRVPSQ 58

Qy         61 AQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASKQKN 120
              |||| :  ||    | ||| | ||| |||| ||| ||::|  |:|  |: |:| || ||
Db         59 AQKLGAGQKPA---KPKQLPAASVPRPVSR-LNNPQKNEQ--PAASGNDSEKEQASLQKT 112

Qy        121 EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRR 180
              |::||||| ||||:||||||||||||||:|||||||||||||| ||||| |||||||
Db        113 EDTKKRQWTLEDFDIGRPLGKGKFGNVYLARERQSKFILALKVLFKTQLEKANVEHQLRR 172

Qy        181 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 EVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITEL 232

Qy        241 ANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEM 300
              |||||||||||||||||||||||||| |||||||||||||||||:||||||||||||
```

```
Db        233 ANALSYCHSKRVIHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTMCGTLDYLPPEM 292

Qy        301 IEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLI 360
              ||||||||||||||||||||||||| |||||:|||||| ||||||||||||||||||||
Db        293 IEGRMHDEKVDLWSLGVLCYEFLVGMPPFEAHTYQETY-RISRVEFTFPDFVTEGARDLI 351

Qy        361 SRLLKHNPSQRPMLREVLEHPWITANSSK-PSNCQNKESASKQS 403
              ||||||| |||  | ||||||| ||||| |:  :||  || |
Db        352 SRLLKHNASQRLTLAEVLEHPWIKANSSKPPTGHTSKEPTSKSS 395


RESULT 3
S52242
protein kinase (EC 2.7.1.-) p46XlEg22 - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
C;Accession: S52242
R;Roghi, C.; le Guellec, R.; Paris, J.; Couturier, A.; Philippe, M.
submitted to the EMBL Data Library, October 1992
A;Description: Eg2, selected by differential screening encodes a new Xenopus protein kin
A;Reference number: S52242
A;Accession: S52242
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-389 <ROG>
A;Cross-references: EMBL:Z17206; NID:g609279; PIDN:CAA78914.1; PID:g609280
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase
F;119-371/Domain: protein kinase homology <KIN>
F;127-135/Region: protein kinase ATP-binding motif

  Query Match             62.1%;  Score 1310.5;  DB 2;  Length 389;
  Best Local Similarity   68.3%;  Pred. No. 7.1e-47;
  Matches  258;  Conservative   44;  Mismatches   65;  Indels   11;  Gaps    4;

Qy         21 GPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQAQK-LVSSHKPVQNQKQKQL 79
              ||||: |:|    |   ||    |||:| |||  ||| :|||| ::|: ||     |  |
Db          4 GPKRIPVSQPPSTQVRPPVTGVSAQRILGPSNVPQRVMMQAQKPVLSNQKPT---AQGLL 60

Qy         80 QATSVPHPVSRPLN-NTQKSKQPLPSAPENNPEEELASKQ------KNEESKKRQWALED 132
              :  :  |  |:|   |  :: |    :   | |:: :: |        || ||:|| |||
Db         61 RPATHGHQTSKPQGPNENRNPQQTSHSSTPNMEKKGSTDQGKTLAVPKEEGKKKQWCLED 120

Qy        133 FEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPN 192
              ||||||||||||||||||::|||||||||||:|||||||||||||||||||||||||||
Db        121 FEIGRPLGKGKFGNVYLARERESKFILALKVLFKSQLEKAGVEHQLRREVEIQSHLRHPN 180

Qy        193 ILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRV 252
              ||||||||||:|||||:||| | ::||||| ::||:||:| || :|| || |||||:|
Db        181 ILRLYGYFHDASRVYLILDYAPGGELFRELQKCTRFDDQRSALYIKQLAEALLYCHSKKV 240

Qy        253 IHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDL 312
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        241 IHRDIKPENLLLGSNGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDL 300

Qy        313 WSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRP 372
              |||||||||||||||| :|:||||:|||:||| :| :|:|  |:||:|:|||||: |
Db        301 WSLGVLCYEFLVGKPPFETDTHQETYRRISKVEFQYPPYVSEEAKDLVSKLLKHNPNHRL 360

Qy        373 MLREVLEHPWITANSSKP 390
               |: |||||||  || :|
Db        361 PLKGVLEHPWIVKNSQQP 378


RESULT 4
S52243
p46Eg265 protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 07-May-1995 #sequence_revision 03-Aug-1995 #text_change 16-Aug-2004
C;Accession: S52243; S34642; I51695
R;Roghi, C.; le Guellec, R.; Paris, J.; Couturier, A.; Philippe, M.
submitted to the EMBL Data Library, October 1992
A;Description: Eg2, selected by differential screening encodes a new Xenopus protein kin
A;Reference number: S52242
A;Accession: S52243
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-407 <ROG>
A;Cross-references: UNIPROT:Q91820; EMBL:Z17207; NID:g609281; PIDN:CAA78915.1; PID:g60928
R;Bouvet, P.; Omilli, F.; Arlot-Bonnemain, Y.; Legagneux, V.; Roghi, C.; Basset, T.; Osbo
submitted to the EMBL Data Library, June 1993
A;Description: Targetted deadenylation of specific mRNAs in Xenopus embryos by a mechanis
A;Reference number: S34642
A;Accession: S34642
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 328-407 <BOU>
A;Cross-references: EMBL:Z24453; NID:g394756; PIDN:CAA80826.1; PID:g394757
R;Bouvet, P.; Omilli, F.; Arlot-Bonnemains, Y.; Legagneux, V.; Roghi, C.; Bassez, T.; Osb
Mol. Cell. Biol. 14, 1893-1900, 1994
A;Title: The deadenylation conferred by the 3' untranslated region of a developmentally c
A;Reference number: I51695; MUID:94158861; PMID:8114721
A;Accession: I51695
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 328-407 <BO2>
A;Cross-references: EMBL:Z24453; NID:g394756; PIDN:CAA80826.1; PID:g394757
C;Superfamily: protein kinase homology
C;Keywords: ATP; phosphotransferase; protein kinase
F;138-390/Domain: protein kinase homology <KIN>
F;146-154/Region: protein kinase ATP-binding motif

  Query Match             61.4%;  Score 1295.5;  DB 2;  Length 407;
  Best Local Similarity   64.9%;  Pred. No. 3e-46;
  Matches  264;  Conservative   38;  Mismatches   72;  Indels   33;  Gaps    6;

Qy          1 MDRS-KENCISGPVKATAPV-GGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVP 58
              |:|: |||     ||   |:  | ||: | |    |   |  :  |||:| |||  |||
Db          1 MERAVKENHKPSNVKIFHPMTEGAKRIPVNQPQSTQFRPPGTAVSAQRILGPSNVPQRVL 60

Qy         59 LQAQK-LVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSAPENNPEEELASK 117
              |||| ::|| ||           | :|    ||     :| :|       ||::   |
Db         61 AQAQKPILSSQKP----------TTQIP---LRPATQGHQSSKPQGPNENRNPQQTSHSS 107

Qy        118 QKN----------------EESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILA 160
              |                || ||:|| |||||||||||||||||||::||||||
Db        108 TPNVEKKGSTDQGKTSAVPKEEGKKKQWCLEDFEIGRPLGKGKFGNVYLARERESKFILA 167

Qy        161 LKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYR 220
              ||||||:|||||||||||||||||||||||||||||||||||:||||||:||| | ::|
Db        168 LKVLFKSQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDASRVYLILDYAPGGELFR 227

Qy        221 ELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVH 280
              |||| ::||:||:| || :|| || |||||:|||||||||||||| ||||||||||||
Db        228 ELQKCTRFDDQRSAMYIKQLAEALLYCHSKKVIHRDIKPENLLLGSNGELKIADFGWSVH 287

Qy        281 APSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKR 340
              |||||||||||||||||||||||| ||||||||||||||||||| :|:||||:|
Db        288 APSSRRTTLCGTLDYLPPEMIEGRMHDETVDLWSLGVLCYEFLVGKPPFETDTHQETYRR 347

Qy        341 ISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANS 387
              ||:||| :| :|:| ||||:|:|||||: |  |: |||||||  ||
Db        348 ISKVEFQYPPYVSEEARDLVSKLLKHNPNHRLPLKGVLEHPWIIKNS 394


RESULT 5
JC4665
protein kinase (EC 2.7.1.37) - mouse
N;Alternate names: serine/threonine protein kinase; STK-1 protein
C;Species: Mus musculus (house mouse)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 17-Mar-1999
C;Accession: JC4665
R;Niwa, H.; Abe, K.; Kunisada, T.; Yamamura, K.
Gene 169, 197-201, 1996
A;Title: Cell-cycle-dependent expression of the STK-1 gene encoding a novel murine putati
A;Reference number: JC4665; MUID:96194801; PMID:8647446
```

A;Accession: JC4665
A;Molecule type: mRNA
A;Residues: 1-345 <NIW>
A;Cross-references: DDBJ:D21099
A;Experimental source: embryonic stem cell
A;Note: The authors translated the codon GCT for residue 328 as Thr
C;Comment: This enzyme regulates a wide array of cellular responses to changing environm
C;Genetics:
A;Gene: stk-1
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: ATP; phosphotransferase; protein kinase
F;80-332/Domain: protein kinase homology <KIN>
F;88-96/Region: protein kinase ATP-binding motif
F;111/Binding site: ATP (Lys) #status predicted
F;205/Active site: Asp #status predicted

Query Match 49.7%; Score 1049; DB 2; Length 345;
Best Local Similarity 58.2%; Pred. No. 2.5e-36;
Matches 206; Conservative 49; Mismatches 75; Indels 24; Gaps 4;

```
Qy         36 PLPVNSGQAQRVLCPSNSSQRVPLQAQKLVSSHKPVQNQKQKQLQATSVPHPVSRPLNNT 95
              | |  |  :|  |  :  |||| |: :   :|   : |                   :|:
Db          9 PWPYGSKTSQSGL--NTLSQRV-LRKEPATTSALALVN-----------------WSNS 47

Qy         96 QKSKQPLPSAPENNPEEELASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQS 155
              | :  |     ||  :   ||:       |: : :::||||||||||||||||||||:|
Db         48 QSTAAPGQKLAENKSQGSTASQ---GSQNKQPFTIDNFEIGRPLGKGKFGNVYLAREKKS 104

Qy        156 KFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPL 215
              :||:|||:|||:|:|| |||||||||:|||:||:|||||:|| ||:|  |:||||||||
Db        105 RFIVALKILFKSQIEKEGVEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPR 164

Qy        216 GTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADF 275
              | :|:||||   |||||||| : ||::||:||| |:|||||||||||||  ||||||||
Db        165 GELYKELQKSRTFDEQRTATIMEELSDALTYCHKKKVIHRDIKPENLLLGLQGELKIADF 224

Qy        276 GWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQ 335
              |||||||| || |:|||||||||||||||:| |||| :|||||| :|| ||||: ::
Db        225 GWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHS 284

Qy        336 ETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK 389
              |||:|| :|:  ||  |  ||:||| :|||||| || | ||  |||: ||| :
Db        285 ETYRRIVKVDLKFPSSVPSGAQDLIFKLLKHNPWQRLPLAEVATHPWVRANSRR 338
```

Search completed: June 20, 2005, 20:07:58
Job time : 37.5274 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 20, 2005, 19:46:50 ; Search time 122.495 Seconds
(without alignments)
1086.128 Million cell updates/sec

Title: US-10-734-126-3
Perfect score: 1809
Sequence: 1 MAQKENSYPWPYGRQTAPSG.....PWVRANSRRVLPPSALQSVA 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 56

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 49%
Maximum Match 100%
Listing first 65000 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1809	100.0	344	2	AAW18083	Aaw18083 Human Aur
2	1809	100.0	344	2	AAY22475	Aay22475 Human AUR
3	1809	100.0	344	4	AAG67615	Aag67615 Amino aci
4	1809	100.0	344	4	AAG67436	Aag67436 Amino aci
5	1809	100.0	344	6	ABR48188	Abr48188 Human bla
6	1809	100.0	344	6	ABU56641	Abu56641 Lung canc
7	1809	100.0	344	6	ABR92152	Abr92152 Human cer
8	1809	100.0	344	7	ADB80537	Adb80537 Ovarian c
9	1809	100.0	344	7	ADN38885	Adn38885 Cancer/an
10	1809	100.0	344	7	ADN95464	Adn95464 Human BEC
11	1809	100.0	344	8	ADK67718	Adk67718 Human mod
12	1809	100.0	344	8	ADQ89834	Adq89834 Antagonis
13	1809	100.0	344	8	ABM81826	Abm81826 Tumour-as
14	1803	99.7	344	7	ABR61580	Abr61580 Human HsA
15	1797.5	99.4	347	2	AAY27052	Aay27052 Human pro
16	1797.5	99.4	347	6	ABU57643	Abu57643 Different
17	1770.5	97.9	343	8	ADM72221	Adm72221 Human TAS
18	1529	84.5	312	7	ADC10236	Adc10236 Human NOV
19	1520.5	84.1	343	2	AAW99783	Aaw99783 Rat AIM-1
20	1201	66.4	275	7	ABR61581	Abr61581 Human HsA
21	1201	66.4	275	8	ADK67719	Adk67719 Human mod
22	1181	65.3	355	5	ABU65047	Abu65047 Human NOV
23	1171.5	64.8	356	8	ADN61745	Adn61745 Human nov
24	1124	62.1	320	8	ADO57331	Ado57331 Kidney de
25	1057.5	58.5	403	2	AAW18084	Aaw18084 Human Aur
26	1057.5	58.5	403	2	AAY22476	Aay22476 Human AUR
27	1057.5	58.5	403	4	AAG67614	Aag67614 Amino aci
28	1057.5	58.5	403	4	AAG67435	Aag67435 Amino aci
29	1057.5	58.5	403	6	ABP97367	Abp97367 Human ser
30	1057.5	58.5	403	6	ABP97366	Abp97366 Human ser
31	1057.5	58.5	403	6	ABR48160	Abr48160 Human bla
32	1057.5	58.5	403	7	ADF61840	Adf61840 Human ser
33	1057.5	58.5	403	7	ADN39888	Adn39888 Cancer/an
34	1057.5	58.5	403	8	ADE15337	Ade15337 Truncated
35	1057.5	58.5	403	8	ADH59556	Adh59556 Monoclona
36	1057.5	58.5	403	8	ADK67720	Adk67720 Human mod
37	1057.5	58.5	403	8	ADM72205	Adm72205 Human TAS
38	1057.5	58.5	403	8	ADQ89832	Adq89832 Antagonis
39	1057.5	58.5	403	8	ADR05174	Adr05174 Human GTP
40	1057.5	58.5	420	5	ABP41958	Abp41958 Human ova
41	1050.5	58.1	319	6	ABP97470	Abp97470 Truncated
42	1050.5	58.1	403	7	ADJ31727	Adj31727 Human mit
43	1045.5	57.8	402	7	ABR61579	Abr61579 Human HsA
44	1045.5	57.8	402	7	ADD89969	Add89969 Human can
45	1043.5	57.7	309	6	ABP97472	Abp97472 Truncated
46	1041.5	57.6	300	6	ABP97473	Abp97473 Truncated
47	966	53.4	403	5	AAO18740	Aao18740 Human NOV
48	940	52.0	294	8	ADJ49242	Adj49242 Oil-assoc
49	937	51.8	294	8	ADJ49602	Adj49602 Oil-assoc
50	934.5	51.7	295	8	ADJ48641	Adj48641 Oil-assoc
51	933.5	51.6	298	8	ADJ48640	Adj48640 Oil-assoc
52	921.5	50.9	282	8	ADJ49236	Adj49236 Oil-assoc
53	918	50.7	368	8	ADJ48378	Adj48378 Maize oil
54	917.5	50.7	418	8	ADJ48815	Adj48815 Oil-assoc
55	908	50.2	298	8	ADJ48546	Adj48546 Oil-assoc
56	906	50.1	432	8	ADJ50416	Adj50416 Oil-assoc

ALIGNMENTS

RESULT 1
AAW18083
ID AAW18083 standard; protein; 344 AA.
XX
AC AAW18083;
XX
DT 07-SEP-1997 (first entry)
XX
DE Human Aurora-1.
XX
KW Aurora-1; AUR-1; signal transduction; protein kinase; tumour; cancer;
KW protein kinase; gene therapy; diagnosis; antibody.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1. .74
FT /label= N-terminal_domain
FT Domain 74. .344
FT /label= Kinase_domain
FT Modified-site 232
FT /label= Phosphorylation
FT /note= "cAMP-dependent protein kinase phosphorylation
FT site conserved in AUR-2 and yeast and Drosophila
FT homologues"
FT Misc-difference 265
FT /note= "deduced residue from some cDNA clones is Gly"
XX
PN WO9722702-A1.
XX
PD 26-JUN-1997.
XX
PF 25-NOV-1996; 96WO-US018859.
XX
PR 18-DEC-1995; 95US-0008809P.
PR 14-AUG-1996; 96US-0023943P.
XX

PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Mossie KG;
XX
DR WPI; 1997-341693/31.
DR N-PSDB; AAT67289.
XX
PT Aurora-1 and Aurora-2 and related genes - useful in tumour gene therapy.
XX
PS Claim 3; Page 82-83; 98pp; English.
XX
CC Novel human proteins, termed Aurora-1 (AAW18083) and Aurora-2 (AAW18084)
CC (AUR-1 and AUR-2), are related serine/threonine kinases with short N-
CC terminal extensions that appear to be involved in cancer and/or signal
CC transduction disorders. Their amino acid sequences were deduced from
CC pancreatic tumour cDNA clones (AAT67289- 90). AUR-1 and AUR-2 appear to
CC regulate nuclear division, with disruption of their signaling resulting
CC in polyploid cells. AUR-1 RNA is broadly expressed in rapidly dividing
CC cells from normal and tumour tissues. AUR polypeptides can be expressed
CC in host cells and used to raise diagnostic antibodies and to screen for
CC compounds that interact with AUR-1 and/or AUR-2
XX
SQ Sequence 344 AA;

Query Match 100.0%; Score 1809; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.8e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 2
AAY22475
ID AAY22475 standard; protein; 344 AA.
XX
AC AAY22475;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human AUR1 protein sequence.
XX
KW AUR1; AUR2; human; AUR modulator; cancer; glioma; medullablastoma;
KW chondrosarcoma; pancreatic tumour; proliferative disease; diagnosis;
KW therapy.
XX
OS Homo sapiens.
XX
PN WO9937788-A2.
XX
PD 29-JUL-1999.
XX
PF 21-JAN-1999; 99WO-US001283.
XX
PR 22-JAN-1998; 98US-00012135.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Mossie K;
XX
DR WPI; 1999-458699/38.
DR N-PSDB; AAX99724.
XX
PT New nucleic acid encoding human AUR1 and 2 polypeptides, used to identify
PT specific modulators for treating cancer or for diagnosis.
XX
PS Claim 11; Page 140-141; 153pp; English.
XX
CC This sequence is the human AUR1 protein of the invention. The AUR1 and
CC AUR2 proteins can be used to identify specific modulators of, and to
CC generate specific antibodies recognising AUR1 and AUR2. The modulators
CC can be used for treating conditions involving abnormal AUR signal
CC transduction, specifically cancer (of colon, breast, kidney, ovary,
CC bladder, head or neck, also glioma, medullablastoma, chondrosarcoma and
CC pancreatic tumours, particularly of colon (specifically), breast or
CC kidney). The modulators can also be used for studying their effects in
CC animal models of proliferative disease. Probes, based on the coding
CC sequences are used, diagnostically, to detect or quantify AUR mRNA, by
CC hybridisation or polymerase chain reaction (PCR). The DNA, optionally
CC mutated, are useful in gene therapy. Ab are used as diagnostic
CC immunoassay reagents for detecting the proteins
XX
SQ Sequence 344 AA;

Query Match 100.0%; Score 1809; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.8e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 3
AAG67615
ID AAG67615 standard; protein; 344 AA.
XX
AC AAG67615;
XX
DT 26-NOV-2001 (first entry)
XX
DE Amino acid sequence of a human protein.
XX
KW Human; protein kinase; protein phosphatase; signal transduction.
XX

OS Homo sapiens.
XX
PN WO200109316-A1.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-JP005061.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 18-OCT-1999; 99US-0159590P.
PR 11-JAN-2000; 2000JP-00118776.
PR 17-FEB-2000; 2000US-0183322P.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
PI Senoo C, Nezu J;
XX
DR WPI; 2001-570286/64.
XX
PT New genes encoding proteins with protein kinase/protein phosphatase
PT activity, useful in the diagnosis and treatment of diseases.
XX
PS Example 4; Page 97-100; 233pp; Japanese.
XX
CC The specification describes human protein kinase/protein phosphatases. It
CC is expected that the protein kinase/protein phosphatase gene participates
CC in signal transduction in cells. The protein kinase/protein phosphatase
CC polypeptides and polynucleotides are useful for developing diagnostics
CC and treatment agents for human and animal diseases. The protein
CC kinase/protein phosphatase polypeptides are useful as target molecules in
CC designing novel drugs. The protein kinase/protein phosphatase
CC polynucleotides are useful as a source of probes and primers, which may
CC be used to isolate homologous sequences. The present sequence represents
CC a human protein, which is used in the course of the invention
XX
SQ Sequence 344 AA;

Query Match 100.0%; Score 1809; DB 4; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.8e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 4
AAG67436
ID AAG67436 standard; protein; 344 AA.
XX
AC AAG67436;
XX
DT 26-NOV-2001 (first entry)
XX
DE Amino acid sequence of a human polypeptide.
XX
KW Human; protein kinase; protein phosphatase; signal transduction;
KW intracellular signalling pathway.
XX
OS Homo sapiens.
XX
PN WO200109345-A1.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-JP005060.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 18-OCT-1999; 99US-0159590P.
PR 11-JAN-2000; 2000JP-00118776.
PR 17-FEB-2000; 2000US-0183322P.
PR 02-MAY-2000; 2000JP-00183767.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
PI Senoo C, Nezu J;
XX
DR WPI; 2001-564736/63.
XX
PT New genes encoding protein kinase and protein phosphatase, useful for
PT identifying modulators which can be used to treat human or animal
PT disorders associated with the expression or function of these enzymes.
XX
PS Example 4; Page 199-201; 336pp; Japanese.
XX
CC The specification describes human protein kinase/protein phosphatases.
CC The polypeptides are expected to participate in signal transduction in
CC cells. The kinase phosphatases are connected with intracellular
CC signalling pathways. Antisense oligonucleotides and compounds identified
CC by screening (agonists or antagonists) can be used to treat human or
CC animal disorders associated with the expression or function of the
CC protein. In addition, the polypeptides may be used as target molecules
CC fir drug development. The present sequence represents a polypeptide, used
CC in the course of the invention
XX
SQ Sequence 344 AA;

Query Match 100.0%; Score 1809; DB 4; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.8e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
```

```
Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              ||||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 5
ABR48188
ID ABR48188 standard; protein; 344 AA.
XX
AC ABR48188;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated protein sequence SEQ ID NO:93.
XX
KW Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003003906-A2.
XX
PD 16-JAN-2003.
XX
PF 03-JUL-2002; 2002WO-US021338.
XX
PR 03-JUL-2001; 2001US-0302814P.
PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Aziz N;
XX
DR WPI; 2003-201532/19.
DR N-PSDB; ACC51000.
XX
PT Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
PS Claim 10; Page 265; 307pp; English.
XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridises to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
SQ Sequence 344 AA;

Query Match 100.0%; Score 1809; DB 6; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.8e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              ||||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 6
ABU56641
ID ABU56641 standard; protein; 344 AA.
XX
AC ABU56641;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polypeptide #234.
XX
KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
PN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
DR N-PSDB; ABX76370.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 27; Page 371; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by

```
CC   administering a modulatory compound identified. The methods are useful
CC   for treating lung cancer, such as small cell lung cancer, non-small cell
CC   lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC   emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC   hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC   bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC   for diagnostic purposes and as targets for screening for therapeutic
CC   compounds that modulate lung cancer, such as antibodies. Sequences
CC   ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC   invention
XX
SQ   Sequence 344 AA;

  Query Match             100.0%;  Score 1809;  DB 6;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-186;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 7
ABR92152
ID   ABR92152 standard; protein; 344 AA.
XX
AC   ABR92152;
XX
DT   10-SEP-2003  (first entry)
XX
DE   Human cervical cancer cell marker protein SEQ ID NO:214.
XX
KW   Human; cervical cancer; cervical cancer marker; cancer therapy;
KW   detection; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2002101075-A2.
XX
PD   19-DEC-2002.
XX
PF   12-JUN-2002; 2002WO-US018638.
XX
PR   13-JUN-2001; 2001US-0298155P.
PR   13-JUN-2001; 2001US-0298159P.
PR   14-NOV-2001; 2001US-0335936P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Schlegel R,  Chen Y,  Zhao X,  Monahan JE,  Kamatkar S;
PI   Gannavarapu M,  Glatt K,  Hoersch S;
XX
DR   WPI; 2003-156967/15.
DR   N-PSDB; ACF12935.
XX
PT   New isolated nucleic acid molecule useful for detecting, characterizing,
PT   preventing and treating human cervical cancers, in various prognostic and
PT   diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
PS   Claim 4; Page 366-367; 386pp; English.
XX
CC   ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
CC   given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC   normal indicates the presence of cervical cancer. Also described: (1) a
CC   vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC   assessing (M1) whether a patient is afflicted with cervical cancer,
CC   comprising comparing the level of expression of a marker in a patient's
CC   sample, and the normal level of expression of the marker in a control non
CC   -cervical cancer sample, where a significant increase in the level of
CC   expression of the marker in the patient's sample relative to that in the
CC   control sample is an indication that the patient is afflicted with
CC   cervical cancer. (I) has cytostatic activity, and can be used in gene
CC   therapy and in vaccines. (I) is useful in detecting, characterising,
CC   preventing and treating human cervical cancers. (I) may also be used in
CC   various prognostic and diagnostic assays, pharmacogenomics and in
CC   monitoring clinical trials
XX
SQ   Sequence 344 AA;

  Query Match             100.0%;  Score 1809;  DB 6;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-186;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 8
ADB80537
ID   ADB80537 standard; protein; 344 AA.
XX
AC   ADB80537;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Ovarian cancer-associated protein #51.
XX
KW   cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW   post-operative chemotherapy; radiation therapy; tumour prognosis;
KW   pre-cancerous lesion detection.
XX
OS   Homo sapiens.
XX
PN   WO2002102235-A2.
XX
```

```
PD   27-DEC-2002.
XX
PF   18-JUN-2002; 2002WO-US019297.
XX
PR   18-JUN-2001; 2001US-0299234P.
PR   27-AUG-2001; 2001US-0315287P.
PR   05-SEP-2001; 2001US-0317544P.
PR   13-NOV-2001; 2001US-0350666P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Mack DH,  Gish KC;
XX
DR   WPI; 2003-167431/16.
DR   N-PSDB; ADB80536.
XX
PT   Detecting an ovarian cancer-associated transcript in a cell from a
PT   patient, comprises contacting a biological sample from the patient with a
PT   polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS   Claim 13; Page 310; 332pp; English.
XX
CC   The invention relates to a method of detecting an ovarian cancer-
CC   associated transcript in a cell from a patient, by contacting a
CC   biological sample from the patient with a polynucleotide that selectively
CC   hybridizes to a sequence at least 80% identical to any of one of 80
CC   nucleic acid sequences given in the specification. The method is useful
CC   in diagnosing ovarian cancer and in identifying and using agents and/or
CC   targets that inhibit ovarian cancer. The nucleic acid molecule,
CC   polypeptide and the antibody may also be used in detecting ovarian
CC   cancers, monitoring and early detection of relapse following treatment,
CC   monitoring response to therapy, selecting patients for post-operative
CC   chemotherapy or radiation therapy, in selecting mode of therapy,
CC   determining tumour prognosis, early detection of pre-cancerous lesions,
CC   and as vaccines. This sequence corresponds to one of the proteins used
CC   for the detection method of the invention.
XX
SQ   Sequence 344 AA;

  Query Match             100.0%;  Score 1809;  DB 7;  Length 344;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-186;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 9
ADN38885
ID   ADN38885 standard; protein; 344 AA.
XX
AC   ADN38885;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:203.
XX
KW   Human; differential expression; cancer; angiogenic disorder;
KW   fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW   inflammatory disease; autoimmune disease;
KW   retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW   detection; diagnosis; prognosis; drug screening; drug targeting;
KW   wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW   vulnerary; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003042661-A2.
XX
PD   22-MAY-2003.
XX
PF   13-NOV-2002; 2002WO-US036810.
XX
PR   13-NOV-2001; 2001US-0350666P.
PR   21-NOV-2001; 2001US-0332464P.
PR   29-NOV-2001; 2001US-0334393P.
PR   03-DEC-2001; 2001US-0335394P.
PR   14-DEC-2001; 2001US-0340376P.
PR   08-JAN-2002; 2002US-0347211P.
PR   10-JAN-2002; 2002US-0347349P.
PR   08-FEB-2002; 2002US-0355250P.
PR   13-FEB-2002; 2002US-0356714P.
PR   20-FEB-2002; 2002US-0359077P.
PR   29-MAR-2002; 2002US-0368809P.
PR   04-APR-2002; 2002US-0370110P.
PR   12-APR-2002; 2002US-0372246P.
PR   05-JUN-2002; 2002US-0386614P.
PR   16-JUL-2002; 2002US-0396839P.
PR   22-JUL-2002; 2002US-0397775P.
PR   22-JUL-2002; 2002US-0397845P.
PR   09-SEP-2002; 2002US-0409450P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D,  Aziz N,  Ginsburg WM,  Gish KC,  Glynne R,  Hevezi PA;
PI   Mack DH,  Murray R,  Watson SR,  Wilson KE,  Zlotnik A;
XX
DR   WPI; 2003-468649/44.
DR   N-PSDB; ADN38884.
XX
PT   Determining the presence or absence of a pathological cell in a patient,
PT   useful for diagnosing, prognosing or treating cancer, comprises detecting
PT   a nucleic acid in a biological sample.
XX
PS   Claim 12; SEQ ID NO 203; 1385pp; English.
XX
CC   The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC   whose expression is upregulated or downregulated in specific cancers or
CC   other diseases such as angiogenic or fibrotic disorders, and to methods
CC   of determining the presence or absence of a pathological cell in a
CC   patient by detecting a nucleic acid at least 80% identical to those of
CC   the invention or by detecting a polypeptide of the invention. The
CC   invention also relates to expression vectors and host cells comprising a
CC   nucleic acid of the invention; antibodies which specifically bind a
CC   polypeptide of the invention; use of such antibodies for drug targeting;
CC   and methods of screening for modulators of activity or expression of the
CC   polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC   antibodies and methods are useful for diagnosing, prognosing and treating
CC   cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC   atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC   neovascularistaion syndromes, scarring and uterine fibroids. They may
CC   also be useful in wound healing and in contraception. The present
CC   sequence represents a polypeptide of the invention.
XX
```

```
SQ   Sequence 344 AA;

  Query Match           100.0%;  Score 1809;  DB 7;  Length 344;
  Best Local Similarity  100.0%;  Pred. No. 8.8e-186;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 10
ADN95464
ID   ADN95464 standard; protein; 344 AA.
XX
AC   ADN95464;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human BEC/LEC-related protein sequence SeqID387.
XX
KW   growth; differentiation; blood endothelial cell; BEC;
KW   lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW   lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
KW   vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;
KW   inflammatory disease; cancer metastasis; lymphatic system; human.
XX
OS   Homo sapiens.
XX
PN   WO2003080640-A1.
XX
PD   02-OCT-2003.
XX
PF   07-MAR-2003; 2003WO-US006900.
XX
PR   07-MAR-2002; 2002US-0363019P.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (LICN ) LICENTIA LTD.
XX
PI   Alitalo K,  Makinen T,  Petrova T,  Saharinen P,  Saharinen J;
XX
DR   WPI; 2003-876899/81.
DR   N-PSDB; ADN95465.
XX
PS   Example 1; SEQ ID NO 387; 176pp; English.
XX
CC   This invention relates to a method of differentially modulating the
CC   growth or differentiation of blood endothelial cells (BEC) or lymphatic
CC   endothelial cells (LEC) comprises contacting endothelial cells with a
CC   composition comprising an agent that differentially modulates blood or
CC   lymphatic endothelial cells. Treating hereditary lymphoedema comprises
CC   identifying a human subject with lymphoedema and with a mutation in at
CC   least one allele of a gene encoding a LEC protein, where the mutation
CC   correlates with lymphoedema in human subjects, and with the proviso that
CC   the LEC protein is not VEGFR-3; and administering to the subject a
CC   composition comprising a lymphatic growth agent selected from VEGF-C or
CC   VEGF-D polypeptides and polynucleotides. The invention may be useful for
CC   the development of compounds with an antiangiogenic, cytostatic,
CC   vasotropic or antiinflammatory activity or for gene therapy. The method
CC   is useful in modulating the growth or differentiation of blood
CC   endothelial cells or lymphatic endothelial cells, in treating hereditary
CC   lymphoedema, in screening for an endothelial cell disorder or
CC   predisposition to the disorder or in monitoring the efficacy or toxicity
CC   of a drug on endothelial cells. The agent is useful in manufacturing a
CC   medicament for the differential modulation of blood vessel endothelial
CC   cell or lymphatic vessel endothelial cell growth or differentiation. The
CC   lymphatic growth agent may also be used in manufacturing a medicament for
CC   the treatment of hereditary lymphoedema resulting from a mutation in a
CC   LEC gene or of other diseases involving the lymphatic vessels, such as
CC   various inflammatory diseases and cancer metastasis via the lymphatic
CC   system. The present sequence is that of a human LEC/BEC differentially
CC   expressed protein which is related to the method of the invention.Note:
CC   This sequence does not appear in the specification but was obtained by
CC   the indexer using the source data given in table 14 of the specification.
XX
SQ   Sequence 344 AA;

  Query Match           100.0%;  Score 1809;  DB 7;  Length 344;
  Best Local Similarity  100.0%;  Pred. No. 8.8e-186;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 11
ADK67718
ID   ADK67718 standard; protein; 344 AA.
XX
AC   ADK67718;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human modifier of AXIN (MAX) polypeptide.
XX
KW   Human; modifier of AXIN; MAX; cytostatic; gene therapy; protein kinase;
KW   enzyme.
XX
OS   Homo sapiens.
XX
PN   WO2004013308-A2.
XX
PD   12-FEB-2004.
XX
```

PF 06-AUG-2003; 2003WO-US024560.
XX
PR 06-AUG-2002; 2002US-0401534P.
PR 16-SEP-2002; 2002US-0411153P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Gendreau SB, Dora EG, Lickteig K, Amundsen CD;
XX
DR WPI; 2004-157122/15.
DR N-PSDB; ADK67711.
XX
PT Identifying a candidate AXIN pathway-modulating agent, useful in
PT diagnosing and treating cancer, comprises providing an assay system
PT comprising a MAX polypeptide or nucleic acid.
XX
PS Example 1; SEQ ID NO 9; 74pp; English.
XX
CC The present sequence is that of a human modifier of AXIN (MAX)
CC polypeptide characterised as a serine/threonine kinase. Genetic screens
CC were designed to identify modifiers of the axin pathway in Caenorhabditis
CC elegans, where a reduction of function pry-1 (axin) mutant was used.
CC Genes causing altered phenotypes were isolated, and their human
CC homologues, termed MAX, were identified, including the present sequence.
CC These MAX polynucleotides and polypeptides are attractive targets for the
CC treatment of pathologies associated with a defective AXIN signalling
CC pathway, such as cancer. Modulation of MAX or its binding partner is
CC useful for understanding the association of the AXIN pathway and its
CC members in normal and disease conditions and for developing diagnostic
CC and therapeutic modalities for AXIN related pathologies. MAX-modulating
CC agents that act by inhibiting or enhancing MAX expression, directly or
CC indirectly, e.g. by affecting MAX function can be identified using
CC methods of the invention. MAX modulating agents are useful in diagnosis,
CC therapy and pharmaceutical development. Preferred MAX modulating agents
CC include antisense and phosphorothioate morpholino oligomers.
XX
SQ Sequence 344 AA;

Query Match 100.0%; Score 1809; DB 8; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.8e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 12
ADQ89834
ID ADQ89834 standard; protein; 344 AA.
XX
AC ADQ89834;
XX
DT 21-OCT-2004 (first entry)
XX
DE Antagonist of cell cycle progression polypeptide #132.
XX
KW Cytostatic; cancer; cell division cycle; mitosis; meiosis;
KW cell cycle progression.
XX
OS Homo sapiens.
XX
PN WO2004063362-A2.
XX
PD 29-JUL-2004.
XX
PF 31-DEC-2003; 2003WO-GB005635.
XX
PR 10-JAN-2003; 2003US-0439123P.
PR 06-MAY-2003; 2003US-0468402P.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Glover D, Bell G, Frenz L, Midgley C;
XX
DR WPI; 2004-544089/52.
DR N-PSDB; ADQ89833.
XX
PT New cell cycle progression genes and proteins for modulating cell cycle
PT progression in cells, for preventing, treating or diagnosing cell
PT proliferative diseases (e.g. cancer) or for identifying modulators of
PT mitosis or meiosis.
XX
PS Claim 2; SEQ ID NO 264; 461pp; English.
XX
CC The present invention relates to a polynucleotide for preventing,
CC treating or diagnosing a disease in an individual. The composition or the
CC polypeptide, polynucleotide or RNA precursor, or antibody is useful for
CC diagnosing, preventing or treating diseases (e.g. cell proliferative
CC diseases such as cancer) in an individual. These may also be used for
CC identifying substances capable of binding to or modulating the function
CC of the polypeptide, capable of affecting the function of the
CC corresponding gene, or capable of inhibiting the cell division cycle or
CC cell cycle progression, preferably mitosis and/or meiosis. The present
CC sequence represents an antagonist of cell cycle progression protein
CC sequence.
XX
SQ Sequence 344 AA;

Query Match 100.0%; Score 1809; DB 8; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.8e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 13
ABM81826
ID ABM81826 standard; protein; 344 AA.
XX
AC ABM81826;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO70812, SEQ:4700.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN40117.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 4700; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 344 AA;

Query Match 100.0%; Score 1809; DB 8; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.8e-186;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
```

RESULT 14
ABR61580
ID ABR61580 standard; protein; 344 AA.
XX
AC ABR61580;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human HsAIRK-2 (Aurora A) protein.
XX
KW RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; HsAIRK-2;
KW Aurora A.
XX
OS Homo sapiens.
XX
PN WO2003087395-A2.
XX
PD 23-OCT-2003.
XX
PF 15-APR-2003; 2003WO-IB002972.
XX
PR 15-APR-2002; 2002US-0372483P.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Garbay C, Gigoux V, Camonis J, L'hoste S, Samson J;
XX
DR WPI; 2003-845340/78.
DR N-PSDB; ACF58069.
XX
PT Identifying anti-tumoral compounds, comprises determining the capacity of
PT a compound to inhibit interaction between RasGAP and Drosophila
PT melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and
PT aurora binding protein.
XX
PS Disclosure; Page 62-63; 0pp; English.
XX
CC The invention relates to identifying a biologically active compound with
CC anti-tumoural properties, where the compound is studied for its capacity
CC to inhibit the interaction between: (a) RasGAP and the Drosophila
CC melanogaster Aurora kinase or an orthologue of the kinase, or its
CC fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding
CC protein and the D. melanogaster Aurora kinase or an orthologue of the
CC kinase, or its fragment. The method is useful for identifying anti-
CC tumoural agents. The present sequence represents a human HsAIRK-2 (Aurora
CC A) protein, an orthologue of the D. melanogaster Aurora kinase
XX
SQ Sequence 344 AA;

Query Match 99.7%; Score 1803; DB 7; Length 344;

```
  Best Local Similarity   99.7%;  Pred. No. 3.9e-185;
  Matches  343;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPMGA 300

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||
Db        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344


RESULT 15
AAY27052
ID   AAY27052 standard; protein; 347 AA.
XX
AC   AAY27052;
XX
DT   08-OCT-1999  (first entry)
XX
DE   Human protein kinase (HPKM)-1 (clone ID 2940).
XX
KW   Human protein kinase molecule; HPKM; human; protein kinase;
KW   phosphate group; cancer; immune disorder.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   60
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   73
FT                   /note= "potential phosphorylation site"
FT   Domain          83. .106
FT                   /note= "potential signature sequence for protein kinase
FT                   catalytic domain"
FT   Modified-site   111
FT                   /note= "potential phosphorylation site"
FT   Modified-site   171
FT                   /note= "potential phosphorylation site"
FT   Modified-site   179
FT                   /note= "potential phosphorylation site"
FT   Domain          199. .211
FT                   /note= "potential signature sequence for protein kinase
FT                   catalytic domain"
FT   Modified-site   230
FT                   /note= "potential phosphorylation site"
FT   Modified-site   235
FT                   /note= "potential phosphorylation site"
FT   Modified-site   280
FT                   /note= "potential phosphorylation site"
FT   Modified-site   282
FT                   /note= "potential N-glycosylation site"
FT   Modified-site   284
FT                   /note= "potential phosphorylation site"
FT   Modified-site   316
FT                   /note= "potential phosphorylation site"
FT   Modified-site   334
FT                   /note= "potential phosphorylation site"
XX
PN   WO9938981-A2.
XX
PD   05-AUG-1999.
XX
PF   12-JAN-1999;   99WO-US000661.
XX
PR   30-JAN-1998;   98US-00016000.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Bandman O,  Hillman JL,  Lal P,  Akerblom IE,  Shah P,  Corley NC;
PI   Guegler KJ;
XX
DR   WPI; 1999-479190/40.
DR   N-PSDB; AAX89850.
XX
PT   New human protein kinase molecules useful for treating or preventing
PT   cancer or an immune disorder.
XX
PS   Claim 1; Page 64-65; 77pp; English.
XX
CC   The invention provides human protein kinase molecules (HPKM) (AAY27052-
CC   57) and nucleic acid sequences (AAX89850-55) encoding the HPKM
CC   polypeptides respectively. The HPKM polypeptides can be produced
CC   recombinantly by standard recombinant methodology. Protein kinases add
CC   phosphate groups to proteins. HPKM polypeptide or an HPKM antagonist is
CC   used to treat or prevent cancer or an immune disorder. The present
CC   sequence represents the amino acid sequence of HPKM-1
XX
SQ   Sequence 347 AA;

  Query Match            99.4%;  Score 1797.5;  DB 2;  Length 347;
  Best Local Similarity   99.1%;  Pred. No. 1.6e-184;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    1;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTAT- 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATV 180

Qy        180 --IMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RAIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 240

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 300

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 347


RESULT 16
ABU57643
ID   ABU57643 standard; protein; 347 AA.
XX
AC   ABU57643;
XX
DT   09-APR-2003  (first entry)
XX
DE   Differentially expressed breast cancer associated protein #30.
XX
KW   Breast cancer; differential gene expression; BC-cDNA;
```

```
KW   breast cancer diagnosis; breast cancer monitoring;
KW   breast cancer treatment; breast cancer staging.
XX
OS   Homo sapiens.
XX
PN   US2002156263-A1.
XX
PD   24-OCT-2002.
XX
PF   04-OCT-2001; 2001US-00974298.
XX
PR   05-OCT-2000; 2000US-0238331P.
XX
PA   (CHEN/) CHEN H.
XX
PI   Chen H;
XX
DR   WPI; 2003-182653/18.
XX
PT   New cDNAs, which are differentially expressed in (metastatic) breast
PT   cancer useful for diagnosing or staging, breast cancer, or for monitoring
PT   the treatment of breast cancer in an individual.
XX
PS   Example; SEQ ID NO 136; 30pp; English.
XX
CC   The invention describes a combination of cDNAs (designated BC-cDNAs),
CC   which are differentially expressed in breast cancer. The combination
CC   includes 152 cDNA sequences, or their complements. The protein encoded by
CC   any of these BC-cDNAs is useful for screening several molecules or
CC   compounds to identify at least one ligand that specifically binds the
CC   protein, producing or preparing polyclonal or monoclonal antibodies, or
CC   purifying antibodies from a sample. The antibodies, which specifically
CC   bind the protein differentially expressed in breast cancer is useful for
CC   detecting the expression of a protein in a sample. The BC-cDNAs are also
CC   useful for diagnosing, monitoring the treatment of, or staging, breast
CC   cancer. This is the amino acid sequence of a differentially expressed
CC   breast cancer associated protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from the US patent office at
CC   seqdata.uspto.gov/sequence.html?DocID=20020156263
XX
SQ   Sequence 347 AA;

  Query Match             99.4%;  Score 1797.5;  DB 6;  Length 347;
  Best Local Similarity   99.1%;  Pred. No. 1.6e-184;
  Matches  344;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    1;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTAT- 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATV 180

Qy        180 --IMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RAIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 240

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 300

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 347


RESULT 17
ADM72221
ID   ADM72221 standard; protein; 343 AA.
XX
AC   ADM72221;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Human TASK112 polypeptide.
XX
KW   TASK; tumour-associated kinase; cytostatic; tumour antigen;
KW   cell proliferative disorder; cancer; transgenic; human.
XX
OS   Homo sapiens.
XX
PN   WO2004024064-A2.
XX
PD   25-MAR-2004.
XX
PF   05-SEP-2003; 2003WO-US027894.
XX
PR   11-SEP-2002; 2002US-0410166P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Desauvage FJ,  Wood WI,  Zhang Z;
XX
DR   WPI; 2004-282985/26.
DR   N-PSDB; ADM72220.
XX
PT   New tumor-associated kinase nucleic acids and polypeptides, useful as
PT   hybridization probes for isolating full length TASK DNA, for generating
PT   transgenic animals, in chromosome identification, or for tissue typing.
XX
PS   Claim 12; SEQ ID NO 26; 163pp; English.
XX
CC   The invention relates to new isolated tumour-associated kinase (TASK)
CC   nucleic acid molecules and encoded polypeptides. Cytostatic. The
CC   antibody, oligopeptide or organic molecule that binds to the TASK
CC   polypeptide are useful for treating a mammal having a tumour comprising
CC   cells expressing the polypeptide. Antagonists of TASK are useful for
CC   treating or preventing a cell proliferative disorder (e.g. cancer)
CC   associated with increased expression or activity of TASK polypeptide. The
CC   TASK polynucleotides and polypeptides may be used as hybridization probes
CC   for isolating full length TASK DNA, for generating transgenic animals, in
CC   chromosome identification, or for tissue typing. The present sequence
CC   represents a human TASK polypeptide.
XX
SQ   Sequence 343 AA;

  Query Match             97.9%;  Score 1770.5;  DB 8;  Length 343;
  Best Local Similarity   98.8%;  Pred. No. 1.3e-181;
  Matches  340;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHATSLRRKTMCGTLDYL 240

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||: |||||||||||||||||||| |||||||||||||||||||||||||||||
Db        241 PPEMIEGRIDNEKVDLWCIGVLCYELLVGN-PFESASHNETYRRIVKVDLKFPASVPTGA 299
```

```
Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              ||||||||||||||||||||||||||||||||||||||||||||
Db        300 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 343
```

RESULT 18
ADC10236
ID ADC10236 standard; protein; 312 AA.
XX
AC ADC10236;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human NOVX polypeptide SEQ ID NO: 258.
XX
KW cytostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
KW antiinflammatory; gene therapy; antisense therapy; thyromimetic; NOVX;
KW pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW inflammatory disorder; chromosome mapping; tissue typing;
KW predictive medicine.
XX
OS Homo sapiens.
XX
PN WO2003000842-A2.
XX
PD 03-JAN-2003.
XX
PF 04-JUN-2002; 2002WO-US017443.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0295573P.
PR 12-JUN-2001; 2001US-0297567P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298528P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299949P.
PR 22-JUN-2001; 2001US-0300177P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301550P.
PR 03-JUL-2001; 2001US-0302951P.
PR 31-JUL-2001; 2001US-0308890P.
PR 14-SEP-2001; 2001US-0322297P.
PR 25-SEP-2001; 2001US-0324669P.
PR 03-DEC-2001; 2001US-0337477P.
PR 14-DEC-2001; 2001US-0341562P.
PR 21-FEB-2002; 2002US-0358656P.
PR 21-FEB-2002; 2002US-0359122P.
PR 22-FEB-2002; 2002US-0358978P.
PR 22-FEB-2002; 2002US-0359034P.
PR 22-FEB-2002; 2002US-0359035P.
PR 22-FEB-2002; 2002US-0359121P.
PR 27-FEB-2002; 2002US-0359964P.
PR 01-MAR-2002; 2002US-0360858P.
PR 12-MAR-2002; 2002US-0363430P.
PR 12-MAR-2002; 2002US-0363676P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-MAY-2002; 2002US-0379444P.
PR 04-JUN-2002; 2002US-00379444.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Agee ML, Anderson DW, Berghs C, Casman SJ, Catterton E;
PI Dipippo VA, Edinger SR, Eisen A, Ellerman K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Herrmann JL, Hjalt T, Ji W, Kekuda R;
PI Khramtsov NV, Li L, Liu X, Malyankar UM, Miller CE, Millet I;
PI Ort T, Padigaru M, Patturajan M, Pena CEA, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G, Spaderna SK;
PI Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsobrook JP;
PI Burgess CE, Lepley DM;
XX
DR WPI; 2003-210149/20.
DR N-PSDB; ADC10235.
XX
PT New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-
PT associated disorders, such as cancer, obesity, diabetes and inflammatory
PT or CNS diseases.
XX
PS Claim 1; SEQ ID NO 258; 772pp; English.
XX
CC The invention relates to novel isolated polypeptides, mature form of the
CC polypeptide, a sequence that is 95% identical to the polypeptide or the
CC polypeptide comprising one or more conservative substitutions. The NOVX
CC polypeptide is useful for treating or preventing a pathology associated
CC with the polypeptide e.g. disorders associated with aberrant expression
CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
CC endocrine, CNS and inflammatory disorders. They can also be used in
CC various detection and screening assays, chromosome mapping, tissue typing
CC and predictive medicine. This sequence corresponds to one of the
CC polypeptides of the invention.
XX
SQ Sequence 312 AA;

Query Match 84.5%; Score 1529; DB 7; Length 312;
Best Local Similarity 86.9%; Pred. No. 1.2e-155;
Matches 299; Conservative 2; Mismatches 11; Indels 32; Gaps 2;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLTSRSNVQPTAAPGQKVMEN 60

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              ||||||||||||||||||||||||                  ||  |:
Db         61 SSGTPDILTRHFTIDDFEIGRPLGK-----------------ALLCLWP---------- 92

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 180
                 |    :  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 ----EASSVSSPSHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATI 148

Qy        181 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 MEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYL 208

Qy        241 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 PPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGA 268

Qy        301 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 344
              ||||||||||||||||||||||||||||||||||||||||||||
Db        269 QDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSVA 312
```

RESULT 19
AAW99783
ID AAW99783 standard; protein; 343 AA.
XX
AC AAW99783;
XX
DT 07-JUN-1999 (first entry)
XX
DE Rat AIM-1.
XX
KW Rat; AIM-1; aurora and IPL-1 like midbody-associated protein kinase;
KW cell cycle regulating protein; serine-threonine kinase;
KW cell proliferation; cancer.
XX
OS Rattus sp.
XX
PN WO9909160-A1.

XX
PD 25-FEB-1999.
XX
PF 17-AUG-1998; 98WO-JP003641.
XX
PR 15-AUG-1997; 97JP-00235371.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Tatsuka M, Terada Y;
XX
DR WPI; 1999-181033/15.
DR N-PSDB; AAX19715.
XX
PT DNA encoding cell cycle-regulating protein, aurora and IPL-1 like midbody
PT -associated protein kinase AIM-1 - useful for treatment of diseases
PT associated with abnormal proliferation of cells e.g. cancers.
XX
PS Claim 1; Page 31-35; 44pp; Japanese.
XX
CC The present sequence represents rat aurora and IPL-1 like midbody-
CC associated protein kinase, designated AIM-1. AIM-1 protein inhibitors
CC obtained can be used as remedies for treatment of diseases associated
CC with abnormal cell proliferation, e.g. cancers. The genes and AIM-1
CC proteins can be applied in the screening of substances with serine-
CC threonine kinase inhibitory activity. Hybridised oligonucleotides and
CC peptide nucleic acids can be used to the inhibit expression of the AIM-1
CC proteins
XX
SQ Sequence 343 AA;

Query Match 84.1%; Score 1520.5; DB 2; Length 343;
Best Local Similarity 84.5%; Pred. No. 1.2e-154;
Matches 290; Conservative 20; Mismatches 30; Indels 3; Gaps 1;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              |||||| |||||| :|: |||:||||||||||     |  ||:||| | || ||||: ||
Db          1 MAQKENVYPWPYGSKTSQSGLNTLPQRVLRKEPAVTPAQALMNRSNSQSTAVPGQKLTEN 60

Qy         61 SSGTP---DILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEG 117
                |         : ||||:|||||||||||||||||||||| ||||||:||||||||||
Db         61 KGATALQGSQSRQPFTIDNFEIGRPLGKGKFGNVYLAREKKSRFIVALKILFKSQIEKEG 120

Qy        118 VEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRT 177
              ||||||||||||||| |||||:||||||::|||||||||||||||||||||| |||||||
Db        121 VEHQLRREIEIQAHLKHPNILQLYNYFYDQQRIYLILEYAPRGELYKELQKSGTFDEQRT 180

Qy        178 ATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTL 237
              |||||||:||||||| |||||||||||||||:|||||||||||||||||||||||||||
Db        181 ATIMEELSDALMYCHKKKVIHRDIKPENLLLGLQGELKIADFGWSVHAPSLRRKTMCGTL 240

Qy        238 DYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVP 297
              |||||||||||||||| ||||||||||||:|||||||| ||:|||||||||||||:|:|
Db        241 DYLPPEMIEGRMHNEMVDLWCIGVLCYELMVGNPPFESPSHSETYRRIVKVDLKFPSSMP 300

Qy        298 TGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              ||:|||||||:||||:|||| ||||||||||||||||||||||
Db        301 LGAKDLISKLLKHNPSQRLPLEQVSAHPWVRANSRRVLPPSAL 343
```

RESULT 20
ABR61581
ID ABR61581 standard; protein; 275 AA.
XX
AC ABR61581;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human HsAIRK-3 protein.
XX
KW RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; HsAIRK-3.
XX
OS Homo sapiens.

XX
PN WO2003087395-A2.
XX
PD 23-OCT-2003.
XX
PF 15-APR-2003; 2003WO-IB002972.
XX
PR 15-APR-2002; 2002US-0372483P.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Garbay C, Gigoux V, Camonis J, L'hoste S, Samson J;
XX
DR WPI; 2003-845340/78.
DR N-PSDB; ACF58070.
XX
PT Identifying anti-tumoral compounds, comprises determining the capacity of
PT a compound to inhibit interaction between RasGAP and Drosophila
PT melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and
PT aurora binding protein.
XX
PS Disclosure; Page 64-66; 0pp; English.
XX
CC The invention relates to identifying a biologically active compound with
CC anti-tumoural properties, where the compound is studied for its capacity
CC to inhibit the interaction between: (a) RasGAP and the Drosophila
CC melanogaster Aurora kinase or an orthologue of the kinase, or its
CC fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding
CC protein and the D. melanogaster Aurora kinase or an orthologue of the
CC kinase, or its fragment. The method is useful for identifying anti-
CC tumoural agents. The present sequence represents a human HsAIRK-3
CC protein, an orthologue of the D. melanogaster Aurora kinase
XX
SQ Sequence 275 AA;

Query Match 66.4%; Score 1201; DB 7; Length 275;
Best Local Similarity 83.0%; Pred. No. 2.4e-120;
Matches 224; Conservative 22; Mismatches 24; Indels 0; Gaps 0;

```
Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              |  |:|||||||||||||||||||| |:|||||||||||||||||:||||||||||
Db          2 RRLTVDDFEIGRPLGKGKFGNVYLARLKESHFIVALKVLFKSQIEKEGLEHQLRREIEIQ 61

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              ||| |||||||||:| ||:||||||||||||||||   ||||||||:||:||||
Db         62 AHLQHPNILRLYNYFHDARRVYLILEYAPRGELYKELQKSEKLDEQRTATIIEEVADALT 121

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||| |||||||||||||| :||:|||||||||| |   ||||||||||||||||||
Db        122 YCHDKKVIHRDIKPENLLLGFRGEVKIADFGWSVHTPLPERKTMCGTLDYLPPEMIEGRT 181

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLR 309
              ::|||||||||||||||||| ||||||||:|||||:|||::|| |:| ||:||||:|||
Db        182 YDEKVDLWCIGVLCYELLVGYPPFESASHSETYRRILKVDVRFPLSMPLGARDLISRLLR 241

Qy        310 HNPSERLPLAQVSAHPWVRANSRRVLPPSA 339
              : | |||||||:  ||||:|:|||||||  |
Db        242 YQPLERLPLAQILKHPWVQAHSRRVLPPCA 271
```

RESULT 21
ADK67719
ID ADK67719 standard; protein; 275 AA.
XX
AC ADK67719;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human modifier of AXIN (MAX) polypeptide.
XX
KW Human; modifier of AXIN; MAX; cytostatic; gene therapy; protein kinase;
KW enzyme.
XX

OS Homo sapiens.
XX
PN WO2004013308-A2.
XX
PD 12-FEB-2004.
XX
PF 06-AUG-2003; 2003WO-US024560.
XX
PR 06-AUG-2002; 2002US-0401534P.
PR 16-SEP-2002; 2002US-0411153P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Gendreau SB, Dora EG, Lickteig K, Amundsen CD;
XX
DR WPI; 2004-157122/15.
DR N-PSDB; ADK67712.
XX
PT Identifying a candidate AXIN pathway-modulating agent, useful in
PT diagnosing and treating cancer, comprises providing an assay system
PT comprising a MAX polypeptide or nucleic acid.
XX
PS Example 1; SEQ ID NO 10; 74pp; English.
XX
CC The present sequence is that of a human modifier of AXIN (MAX)
CC polypeptide characterised as a serine/threonine kinase (aurora/PIL1-
CC like). Genetic screens were designed to identify modifiers of the axin
CC pathway in Caenorhabditis elegans, where a reduction of function pry-1
CC (axin) mutant was used. Genes causing altered phenotypes were isolated,
CC and their human homologues, termed MAX, were identified, including the
CC present sequence. These MAX polynucleotides and polypeptides are
CC attractive targets for the treatment of pathologies associated with a
CC defective AXIN signalling pathway, such as cancer. Modulation of MAX or
CC its binding partner is useful for understanding the association of the
CC AXIN pathway and its members in normal and disease conditions and for
CC developing diagnostic and therapeutic modalities for AXIN related
CC pathologies. MAX-modulating agents that act by inhibiting or enhancing
CC MAX expression, directly or indirectly, e.g. by affecting MAX function
CC can be identified using methods of the invention. MAX modulating agents
CC are useful in diagnosis, therapy and pharmaceutical development.
CC Preferred MAX modulating agents include antisense and phosphorothioate
CC morpholino oligomers.
XX
SQ Sequence 275 AA;

Query Match 66.4%; Score 1201; DB 8; Length 275;
Best Local Similarity 83.0%; Pred. No. 2.4e-120;
Matches 224; Conservative 22; Mismatches 24; Indels 0; Gaps 0;

```
Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              |  |:|||||||||||||||||||| |:||||||||||||||||||:||||||||||
Db          2 RRLTVDDFEIGRPLGKGKFGNVYLARLKESHFIVALKVLFKSQIEKEGLEHQLRREIEIQ 61

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              ||| |||||||||||:| ||:|||||||||||||||||   ||||||||:||:||||
Db         62 AHLQHPNILRLYNYFHDARRVYLILEYAPRGELYKELQKSEKLDEQRTATIIEEVADALT 121

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||| |||||||||||||| :||:|||||||||| |   ||||||||||||||||||||
Db        122 YCHDKKVIHRDIKPENLLLGFRGEVKIADFGWSVHTPLPERKTMCGTLDYLPPEMIEGRT 181

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLR 309
              ::|||||||||||||||| ||||||||:||||||:|||::|| |:| ||:||||:|||
Db        182 YDEKVDLWCIGVLCYELLVGYPPFESASHSETYRRILKVDVRFPLSMPLGARDLISRLLR 241

Qy        310 HNPSERLPLAQVSAHPWVRANSRRVLPPSA 339
              : | |||||||:  ||||:|:||||||| |
Db        242 YQPLERLPLAQILKHPWVQAHSRRVLPPCA 271
```

RESULT 22
ABU65047
ID ABU65047 standard; protein; 355 AA.
XX
AC ABU65047;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV5a protein.
XX
KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human.
XX
OS Homo sapiens.
XX
PN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US006908.
XX
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0277338P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.

PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX
DR WPI; 2002-723332/78.
DR N-PSDB; ABX97014.
XX
PT NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
PS Claim 1; Page 106; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABU65041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
SQ Sequence 355 AA;

Query Match 65.3%; Score 1181; DB 5; Length 355;
Best Local Similarity 70.8%; Pred. No. 5.2e-118;
Matches 247; Conservative 27; Mismatches 61; Indels 14; Gaps 6;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              |||||||||||||:||||:|||||  ||| : | | :|  | | :: || |||| :|:||
Db          1 MAQKENSYPWPYGKQTAPAGLSTLLPRVLPRIP-TEAARELPSCADPQPAAAPGHEVVEN 59

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              | |   |||| | :|| | |||||| || :|||||:| ||||||||  |||||| |||||
Db         60 SCGKRSILTRPFLVDDLETGRPLGKDKFVHVYLARKKTSHFIVALKA-FKSQIE-EGVEH 117

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPR----GELYKELQKSCTFDEQR 176
              |:||::||||   ||||| ||||||| |:|| ||||||      |||:||:|| |||::
Db        118 QMRRQMEIQAPFQHPNILSLYNYFYDLRKIYWILEYAPATPTPEELYQELRKSRTFDKKP 177

Qy        177 TATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKT--MC 234
              ||||  |:||||||||||||  ||:||:||| ||:||||:|||  ||||||||||  ||
Db        178 TATITGEVADALMYCHGKKVTPRDMKPDNLLSGLEGELKVADFGCPVHAPSLRRKTRQMC 237

Qy        235 GTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPA 294
              |||||| || |||| | |||||| || | || |||||      :||| | ||||| ||||
Db        238 GTLDYLSPETIEGRAHTEKVDLWYIGALGYEPLVGNP-----THNEAYGRIVKVALKFPL 292

Qy        295 SVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
                |   |||||||||||||||||||||||| : |:||||||||| |||
Db        293 LCPGEPQDLISKLLRHNPSERLPLAQVSAHPGILAHSRRVLPPSAHQSV 341
```

RESULT 23
ADN61745
ID ADN61745 standard; protein; 356 AA.
XX
AC ADN61745;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human novel protein NOV5a.
XX
KW Human; NOVX; diabetes; obesity; infectious disease; anorexia;
KW cancer-associated cachexia; cancer; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; chronic disease; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 47
FT /note= "May be Leu as a result of a single nucleotide
FT polymorphism"
XX
PN US2004043382-A1.
XX
PD 04-MAR-2004.
XX
PF 07-MAR-2002; 2002US-00092900.
XX
PR 08-MAR-2001; 2001US-0274191P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281444P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.

PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
XX
PA (PADI/) PADIGARU M.
PA (SPYT/) SPYTEK K A.
PA (SHEN/) SHENOY S G.
PA (TAUP/) TAUPIER R J.
PA (PENA/) PENA C E A.
PA (LILL/) LI L.
PA (ZERH/) ZERHUSEN B D.
PA (GUSE/) GUSEV V Y.
PA (JIWW/) JI W.
PA (GORM/) GORMAN L.
PA (MILL/) MILLER C E.
PA (KEKU/) KEKUDA R.
PA (PATT/) PATTURAJAN M.
PA (GANG/) GANGOLLI E A.
PA (VERN/) VERNET C A M.
PA (GUOX/) GUO X S.
PA (TCHE/) TCHERNEV V T.
PA (FERN/) FERNANDES E R.
PA (CASM/) CASMAN S J.
PA (MALY/) MALYANKAR U M.
PA (GERL/) GERLACH V.
PA (LIUY/) LIU Y.
PA (ANDE/) ANDERSON D W.
PA (SPAD/) SPADERNA S K.
PA (CATT/) CATTERTON E.
PA (LEIT/) LEITE M W.
PA (ZHON/) ZHONG H.
PA (ALSO/) ALSOBROOK J P.
PA (LEPL/) LEPLEY D M.
PA (RIEG/) RIEGER D K.
PA (BURG/) BURGESS C E.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev VY, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli EA, Vernet CAM, Guo XS, Tchernev VT;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
PI Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;
PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
DR WPI; 2004-225693/21.
DR N-PSDB; ADN61744.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,
PT infection or obesity, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; SEQ ID NO 14; 786pp; English.
XX
CC The invention relates to an isolated polypeptide (designated NOVX, or
CC NOV1-NOV127) comprising a sequence selected from 178 fully defined amino
CC acid sequences (and their mature forms, variants and fragments). Also
CC included are an isolated nucleic acid molecule encoding NOVx, a vector
CC comprising the nucleic acid, a cell comprising the vector, methods for
CC determining the presence or amount of the polypeptide or the nucleic acid
CC molecule in a sample, methods for determining the presence of or
CC predisposition to a disease associated with altered levels of expression
CC of the above polypeptide or nucleic acid molecule in a first mammalian
CC subject, a method for identifying an agent that binds to the above
CC polypeptide, a method for identifying a potential therapeutic agent for
CC use in the treatment of a pathology that is related to aberrant
CC expression or physiological interactions of the polypeptide, a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide and a method for modulating
CC the activity of the polypeptide cited above. The composition and methods
CC are useful for diagnosing, preventing or treating diseases such as
CC diabetes, obesity, infectious diseases, anorexia, cancer-associated
CC cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or
CC Parkinson's disease, immune disorders, haematopoietic disorders,
CC dyslipidaemias, and other chronic diseases. These may also be used in
CC chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The polypeptides are also useful as vaccines. The
CC present sequence represents a NOVX protein of the invention.
XX
SQ Sequence 356 AA;

Query Match 64.8%; Score 1171.5; DB 8; Length 356;
Best Local Similarity 70.6%; Pred. No. 5.5e-117;
Matches 247; Conservative 27; Mismatches 61; Indels 15; Gaps 7;

```
Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMEN 60
              |||||||||||||:||||:|||||  ||| : | | :|  | | :: || |||| :|:||
Db          1 MAQKENSYPWPYGKQTAPAGLSTLLPRVLPRIP-TEAARELPSCADPQPAAAPGHEVVEN 59

Qy         61 SSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEH 120
              | |   |||| | :|| | |||||| || :|||||:| ||||||||  |||||| |||||
Db         60 SCGKRSILTRPFLVDDLETGRPLGKDKFVHVYLARKKTSHFIVALKA-FKSQIE-EGVEH 117

Qy        121 QLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPR----GELYKELQKSCTFDEQR 176
              |:||::|||   ||||| ||||||| |:|| ||||||      |||:||:|| |||::
Db        118 QMRRQMEIQAPFQHPNILSLYNYFYDLRKIYWILEYAPATPTPEELYQELRKSRTFDKKP 177

Qy        177 TATIMEELADALMYCHGKKVIHRDIKPENLLL-GLKGELKIADFGWSVHAPSLRRKT--M 233
              ||||  |:|||||||||||  ||:||:| || ||:||||:||||  ||||||||||  |
Db        178 TATITGEVADALMYCHGKKVTPRDMKPDNTLLSGLEGELKVADFGCPVHAPSLRRKTRQM 237

Qy        234 CGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFP 293
              ||||||| || |||| | |||||| || | || |||||     :||| | ||||| ||||
Db        238 CGTLDYLSPETIEGRAHTEKVDLWYIGALGYEPLVGNP-----THNEAYGRIVKVALKFP 292

Qy        294 ASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
                  |   |||||||||||||||||||||||| : |:||||||||| |||
Db        293 LLCPGEPQDLISKLLRHNPSERLPLAQVSAHPGILAHSRRVLPPSAHQSV 342
```

RESULT 24
ADO57331
ID ADO57331 standard; protein; 320 AA.
XX
AC ADO57331;
XX
DT 15-JUL-2004 (first entry)
XX
DE Kidney development associated protein seqid 98.
XX
KW nephrotropic; cardiant; antiinfertility; cerebroprotective;
KW neuroprotective; muscular; cytostatic; osteopathic; gene therapy;
KW kidney development; kidney disorder; developmental disorder;
KW circulatory disorder; hearing disorder; heart defect; infertility;
KW stroke; mental retardation; muscle defect; proliferative disorder;
KW bone defect; bone disorder; zebrafish.
XX
OS Danio rerio.
XX
PN US2004068763-A1.
XX
PD 08-APR-2004.
XX
PF 28-MAR-2003; 2003US-00403571.
XX

```
PR   29-MAR-2002; 2002US-0368760P.
XX
PA   (HOPK/) HOPKINS N.
PA   (GOLL/) GOLLING G.
PA   (AMST/) AMSTERDAM A.
PA   (SUNZ/) SUN Z.
XX
PI   Hopkins N,  Golling G,  Amsterdam A,  Sun Z;
XX
DR   WPI; 2004-304692/28.
DR   N-PSDB; ADO57330.
XX
PT   New 459 nucleic acids and encoded polypeptides, useful for diagnosing,
PT   treating or preventing a kidney disorder in an organism, or in screening
PT   for compounds that modulate the development of an organism.
XX
PS   Disclosure; SEQ ID NO 98; 347pp; English.
XX
CC   The invention describes an isolated nucleic acid molecule (I) comprising
CC   a sequence having at least 75% sequence identity to the 459 nucleic acid
CC   sequence of 2808 base pairs (SEQ ID NO: 59) given in the specification,
CC   over at least 600 contiguous base pairs, where the nucleic acid functions
CC   in kidney development. (I) is useful for treating or preventing a kidney
CC   disorder in an organism, where the nucleic acid elicits an alteration in
CC   expression of a 459 nucleic acid sequence in the organism and
CC   subsequently treats or prevents a kidney disorder. The nucleic acid may
CC   also be used in diagnosing, preventing and treating a variety of
CC   mammalian diseases and developmental disorders (e.g. circulatory
CC   disorders, hearing disorders, heart defect, infertility, stroke, mental
CC   retardation, muscle defects, proliferative disorders, or bone defects or
CC   disorders) as well as in screening for compounds that modulate the
CC   development of an organism as a whole or of specific tissues or organs
CC   within that organism. This is the amino acid sequence of a kidney
CC   development associated protein.
XX
SQ   Sequence 320 AA;

  Query Match             62.1%;  Score 1124;  DB 8;  Length 320;
  Best Local Similarity   65.7%;  Pred. No. 6.2e-112;
  Matches  224;  Conservative   33;  Mismatches   52;  Indels   32;  Gaps    4;

Qy          1 MAQKENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQP----TAAPGQK 56
              |  |||  |                    || :    |||| |   |  : |     : ||:
Db          1 MQNKENREP----------------RVQQ----TPSAGVGPLRVEMNPDTHAVSGPGRV 39

Qy         57 VMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKE 116
               ::::|      :  :||||:||||||||||||||||:|   ::|||||||||: ||
Db         40 PVKSNS-------KVLSIDDFDIGRPLGKGKFGNVYLARERKLKVVIALKVLFKSQMVKE 92

Qy        117 GVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQR 176
              |||||||||||||:|| |||||| ||||:|  |::||||||||||:|||||:   ||:||
Db         93 GVEHQLRREIEIQSHLRHPNILRFYNYFHDDTRVFLILEYAPRGEMYKELQRYGHFDDQR 152

Qy        177 TATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGT 236
              ||| |||::||| ||| |||||||||||||||| :|||||||||||||||||:|||||
Db        153 TATYMEEVSDALQYCHEKKVIHRDIKPENLLLGYRGELKIADFGWSVHAPSLRRRTMCGT 212

Qy        237 LDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASV 296
              ||||||||||  |:|||||| ||||||| ||||||||:  | |||:|| ||||:||  |
Db        213 LDYLPPEMIEGHSHDEKVDLWSIGVLCYECLVGNPPFETRQHAETYKRITKVDLQFPKLV 272

Qy        297 PTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPP 337
              ||:||||||||:|| ||||  |  |  |:|||||||||
Db        273 SEGARDLISKLLRHSPSMRLPLRSVMEHRGVKANSRRVLPP 313


RESULT 25
AAW18084
ID   AAW18084 standard; protein; 403 AA.
XX
AC   AAW18084;
XX
DT   07-SEP-1997  (first entry)
XX
DE   Human Aurora-2.
XX
KW   Aurora-2; AUR-2; signal transduction; protein kinase; tumour; cancer;
KW   protein kinase; gene therapy; diagnosis; antibody.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .130
FT                   /label= N-terminal_domain
FT   Misc-difference 50
FT                   /note= "deduced residue from some cDNA clones is Leu"
FT   Misc-difference 50
FT                   /note= "deduced residue from some cDNA clones is Leu"
FT   Misc-difference 57
FT                   /note= "deduced residue from some cDNA clones is Ile"
FT   Misc-difference 57
FT                   /note= "deduced residue from some cDNA clones is Ile"
FT   Domain          131. .403
FT                   /label= Kinase_domain
FT   Modified-site   288
FT                   /label= Phosphorylation
FT                   /note= "cAMP-dependent protein kinase phosphorylation
FT                   site conserved in AUR-2 and yeast and Drosophila
FT                   homologues"
FT   Modified-site   334
FT                   /label= Phosphorylation
FT                   /note= "tyrosine phosphorylation consensus site conserved
FT                   in Drosophila auroura but not in AUR-1 or yeast IPL1"
FT   Modified-site   342
FT                   /label= Phosphorylation
FT                   /note= "cAMP-dependent protein kinase phosphorylation
FT                   site conserved in AUR-2 and yeast and Drosophila
FT                   homologues"
XX
PN   WO9722702-A1.
XX
PD   26-JUN-1997.
XX
PF   25-NOV-1996;   96WO-US018859.
XX
PR   18-DEC-1995;   95US-0008809P.
PR   14-AUG-1996;   96US-0023943P.
XX
PA   (SUGE-) SUGEN INC.
XX
PI   Plowman GD,  Mossie KG;
XX
DR   WPI; 1997-341693/31.
DR   N-PSDB; AAT67290.
XX
PT   Aurora-1 and Aurora-2 and related genes - useful in tumour gene therapy.
XX
PS   Claim 3; Page 83-85; 98pp; English.
XX
CC   Novel human proteins, termed Aurora-1 (AAW18083) and Aurora-2 (AAW18084)
CC   (AUR-1 and AUR-2), are related serine/threonine kinases with short N-
CC   terminal extensions that appear to be involved in cancer and/or signal
CC   transduction disorders. Their amino acid sequences were deduced from
CC   pancreatic tumour cDNA clones (AAT67289- 90). AUR-1 and AUR-2 appear to
CC   regulate nuclear division, with disruption of their signaling resulting
CC   in polyploid cells. AUR-2 RNA is low or absent in most normal tissues,
CC   and abundant in a subset of tumour-derived cell lines, partic. those of
CC   colorectal origin. AUR polypeptides can be expressed in host cells and
CC   used to raise diagnostic antibodies and to screen for compounds that
CC   interact with AUR-1 and/or AUR-2
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 2;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;
```

```
Qy          22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
               :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db          82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy          74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
               ::|||||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:|||:||
Db         130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy         134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
               |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db         190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy         194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
               |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db         250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy         254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
               |||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||:|||
Db         310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy         314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
               :|  | :|  |||: ||| :   ||  |:
Db         370 QRPMLREVLEHPWITANSSK---PSNCQN 395
```

RESULT 26
AAY22476
ID AAY22476 standard; protein; 403 AA.
XX
AC AAY22476;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human AUR2 protein sequence.
XX
KW AUR1; AUR2; human; AUR modulator; cancer; glioma; medullablastoma;
KW chondrosarcoma; pancreatic tumour; proliferative disease; diagnosis;
KW therapy.
XX
OS Homo sapiens.
XX
PN WO9937788-A2.
XX
PD 29-JUL-1999.
XX
PF 21-JAN-1999; 99WO-US001283.
XX
PR 22-JAN-1998; 98US-00012135.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Mossie K;
XX
DR WPI; 1999-458699/38.
DR N-PSDB; AAX99725.
XX
PT New nucleic acid encoding human AUR1 and 2 polypeptides, used to identify
PT specific modulators for treating cancer or for diagnosis.
XX
PS Claim 11; Page 141-142; 153pp; English.
XX
CC This sequence is the human AUR2 protein of the invention. The AUR1 and
CC AUR2 proteins can be used to identify specific modulators of, and to
CC generate specific antibodies recognising AUR1 and AUR2. The modulators
CC can be used for treating conditions involving abnormal AUR signal
CC transduction, specifically cancer (of colon, breast, kidney, ovary,
CC bladder, head or neck, also glioma, medullablastoma, chondrosarcoma and
CC pancreatic tumours, particularly of colon (specifically), breast or
CC kidney). The modulators can also be used for studying their effects in
CC animal models of proliferative disease. Probes, based on the coding
CC sequences are used, diagnostically, to detect or quantify AUR mRNA, by
CC hybridisation or polymerase chain reaction (PCR). The DNA, optionally
CC mutated, are useful in gene therapy. Ab are used as diagnostic
CC immunoassay reagents for detecting the proteins
XX
SQ Sequence 403 AA;

Query Match 58.5%; Score 1057.5; DB 2; Length 403;
Best Local Similarity 63.2%; Pred. No. 1.3e-104;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy          22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
               :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db          82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy          74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
               ::|||||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:|||:||
Db         130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy         134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
               |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db         190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy         194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
               |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db         250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy         254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
               |||| :|||||| ||| |||||: :: |||:|| :|:  ||  |  ||:||||:||:|||
Db         310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy         314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
               :|  | :|  |||: ||| :   ||  |:
Db         370 QRPMLREVLEHPWITANSSK---PSNCQN 395
```

RESULT 27
AAG67614
ID AAG67614 standard; protein; 403 AA.
XX
AC AAG67614;
XX
DT 26-NOV-2001 (first entry)
XX
DE Amino acid sequence of a human protein.
XX
KW Human; protein kinase; protein phosphatase; signal transduction.
XX
OS Homo sapiens.
XX
PN WO200109316-A1.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-JP005061.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 18-OCT-1999; 99US-0159590P.
PR 11-JAN-2000; 2000JP-00118776.
PR 17-FEB-2000; 2000US-0183322P.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Funahashi S;
PI Senoo C, Nezu J;
XX
DR WPI; 2001-570286/64.
XX
PT New genes encoding proteins with protein kinase/protein phosphatase
PT activity, useful in the diagnosis and treatment of diseases.
XX
PS Example 4; Page 94-97; 233pp; Japanese.

```
XX
CC   The specification describes human protein kinase/protein phosphatases. It
CC   is expected that the protein kinase/protein phosphatase gene participates
CC   in signal transduction in cells. The protein kinase/protein phosphatase
CC   polypeptides and polynucleotides are useful for developing diagnostics
CC   and treatment agents for human and animal diseases. The protein
CC   kinase/protein phosphatase polypeptides are useful as target molecules in
CC   designing novel drugs. The protein kinase/protein phosphatase
CC   polynucleotides are useful as a source of probes and primers, which may
CC   be used to isolate homologous sequences. The present sequence represents
CC   a human protein, which is used in the course of the invention
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 4;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||   ||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 28
AAG67435
ID   AAG67435 standard; protein; 403 AA.
XX
AC   AAG67435;
XX
DT   26-NOV-2001  (first entry)
XX
DE   Amino acid sequence of a human polypeptide.
XX
KW   Human; protein kinase; protein phosphatase; signal transduction;
KW   intracellular signalling pathway.
XX
OS   Homo sapiens.
XX
PN   WO200109345-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-JP005060.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   18-OCT-1999;   99US-0159590P.
PR   11-JAN-2000; 2000JP-00118776.
PR   17-FEB-2000; 2000US-0183322P.
PR   02-MAY-2000; 2000JP-00183767.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T,  Funahashi S;
PI   Senoo C,  Nezu J;
XX
DR   WPI; 2001-564736/63.
XX
PT   New genes encoding protein kinase and protein phosphatase, useful for
PT   identifying modulators which can be used to treat human or animal
PT   disorders associated with the expression or function of these enzymes.
XX
PS   Example 4; Page 196-199; 336pp; Japanese.
XX
CC   The specification describes human protein kinase/protein phosphatases.
CC   The polypeptides are expected to participate in signal transduction in
CC   cells. The kinase phosphatases are connected with intracellular
CC   signalling pathways. Antisense oligonucleotides and compounds identified
CC   by screening (agonists or antagonists) can be used to treat human or
CC   animal disorders associated with the expression or function of the
CC   protein. In addition, the polypeptides may be used as target molecules
CC   fir drug development. The present sequence represents a polypeptide, used
CC   in the course of the invention
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 4;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||   ||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 29
ABP97367
ID   ABP97367 standard; protein; 403 AA.
XX
AC   ABP97367;
XX
DT   10-MAY-2003  (first entry)
XX
DE   Human serine/threonine kinase 15 (STK15), Ile31 variant.
XX
KW   Human; serine/threonine kinase 15; STK15; STK6; Aurora2; cell cycle;
KW   chromosome 20; centrosome-associated kinase; cancer susceptibility;
KW   single nucleotide polymorphism; SNP; genetic diagnosis; prognosis;
KW   detection; diagnosis; cancer; malignant astrocytoma; glioblastoma;
KW   medulloblastoma; gastric cancer; colorectal cancer; colorectal adenoma;
KW   acute myelogenous leukaemia; lung cancer; renal cancer; leukaemia;
KW   breast cancer; prostate cancer; endometrial cancer; neuroblastoma;
KW   enzyme.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 31
FT                   /note= "This residue is replaced by Phe in a polymorphic
FT                   variant"
FT   Misc-difference 57
FT                   /note= "Encoded by ATT"
XX
PN   WO2003012046-A2.
XX
PD   13-FEB-2003.
XX
PF   29-JUL-2002; 2002WO-US024115.
XX
PR   27-JUL-2001; 2001US-0308911P.
PR   28-NOV-2001; 2001US-0334146P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Toland AE,  Balmain A;
XX
DR   WPI; 2003-239517/23.
DR   N-PSDB; ABZ97367.
XX
PT   Determining cancer susceptibility in a human subject comprises
PT   identifying in a nucleic acid sample from the subject, a nucleotide
PT   occurrence of a single polynucleotide polymorphism (SNP) of the STK15
PT   gene.
XX
PS   Claim 63; Page; 92pp; English.
XX
CC   The invention relates to a method for determining cancer susceptibility
CC   in a human patient. The method involves determining the identity of the
CC   nucleotide at position 457 of the serine/threonine kinase 15 (STK15) DNA
CC   (ABZ75005). This site is a T/A single nucleotide polymorphism (SNP) in
CC   the coding region of the DNA, resulting in either a Phe or Ile residue at
CC   position 31 in the corresponding STK15 protein (ABP97366). The A457
CC   (Ile31) allele (see ABZ75006, ABP97367) is associated with an increased
CC   cancer susceptibility. STK15 (also known as STK6 and Aurora2) is a
CC   centrosome-associated kinase that is highly expressed at the G2 and M
CC   phase of the cell cycle, and its gene is located on chromosome 20. The
CC   method of the invention are useful for determining cancer susceptibility
CC   and for prognosing, detecting and/or diagnosing cancers such as malignant
CC   astrocytoma, glioblastoma, medulloblastoma, gastric cancer, colorectal
CC   cancer, colorectal adenoma, acute myelogenous leukemia, lung cancer,
CC   renal cancer, leukaemia, breast cancer, prostate cancer, endometrial
CC   cancer and neuroblastoma. This sequence represents the Ile31 variant of
CC   human STK15 which is associated with increased cancer susceptibility.
CC   Note: The present sequence is not shown in the specification, but is
CC   derived from the human STK15 sequence shown in the sequence listing
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 6;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 30
ABP97366
ID   ABP97366 standard; protein; 403 AA.
XX
AC   ABP97366;
XX
DT   10-MAY-2003  (first entry)
XX
DE   Human serine/threonine kinase 15 (STK15).
XX
KW   Human; serine/threonine kinase 15; STK15; STK6; Aurora2; cell cycle;
KW   chromosome 20; centrosome-associated kinase; cancer susceptibility;
KW   single nucleotide polymorphism; SNP; genetic diagnosis; prognosis;
KW   detection; diagnosis; cancer; malignant astrocytoma; glioblastoma;
KW   medulloblastoma; gastric cancer; colorectal cancer; colorectal adenoma;
KW   acute myelogenous leukaemia; lung cancer; renal cancer; leukaemia;
KW   breast cancer; prostate cancer; endometrial cancer; neuroblastoma;
KW   enzyme.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 31
FT                   /label= Phe, Ile
FT                   /note= "Encoded by WTT. This position represents a
FT                   polymorphic site"
FT   Misc-difference 57
FT                   /note= "Encoded by ATT"
XX
PN   WO2003012046-A2.
XX
PD   13-FEB-2003.
XX
PF   29-JUL-2002; 2002WO-US024115.
XX
PR   27-JUL-2001; 2001US-0308911P.
PR   28-NOV-2001; 2001US-0334146P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Toland AE,  Balmain A;
XX
DR   WPI; 2003-239517/23.
DR   N-PSDB; ABZ75005.
XX
PT   Determining cancer susceptibility in a human subject comprises
PT   identifying in a nucleic acid sample from the subject, a nucleotide
PT   occurrence of a single polynucleotide polymorphism (SNP) of the STK15
PT   gene.
XX
PS   Claim 63; Page 85-86; 92pp; English.
XX
CC   The invention relates to a method for determining cancer susceptibility
CC   in a human patient. The method involves determining the identity of the
CC   nucleotide at position 457 of the serine/threonine kinase 15 (STK15) DNA
CC   (ABZ75005). This site is a T/A single nucleotide polymorphism (SNP) in
CC   the coding region of the DNA, resulting in either a Phe or Ile residue at
CC   position 31 in the corresponding STK15 protein (ABP97366). The A457
CC   (Ile31) allele (see ABZ75006, ABP97367) is associated with an increased
CC   cancer susceptibility. STK15 (also known as STK6 and Aurora2) is a
CC   centrosome-associated kinase that is highly expressed at the G2 and M
CC   phase of the cell cycle, and its gene is located on chromosome 20. The
CC   method of the invention are useful for determining cancer susceptibility
CC   and for prognosing, detecting and/or diagnosing cancers such as malignant
CC   astrocytoma, glioblastoma, medulloblastoma, gastric cancer, colorectal
```

```
CC   cancer, colorectal adenoma, acute myelogenous leukemia, lung cancer,
CC   renal cancer, leukaemia, breast cancer, prostate cancer, endometrial
CC   cancer and neuroblastoma. This sequence represents human STK15
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 6;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||     :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:||||   ||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 31
ABR48160
ID   ABR48160 standard; protein; 403 AA.
XX
AC   ABR48160;
XX
DT   12-JUN-2003  (first entry)
XX
DE   Human bladder cancer associated protein sequence SEQ ID NO:35.
XX
KW   Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003003906-A2.
XX
PD   16-JAN-2003.
XX
PF   03-JUL-2002; 2002WO-US021338.
XX
PR   03-JUL-2001; 2001US-0302814P.
PR   03-AUG-2001; 2001US-0310099P.
PR   08-NOV-2001; 2001US-0343705P.
PR   13-NOV-2001; 2001US-0350666P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Mack DH,  Aziz N;
XX
DR   WPI; 2003-201532/19.
DR   N-PSDB; ACC50970.
XX
PT   Detecting a bladder cancer-associated transcript in a cell from a
PT   patient, comprises contacting a biological sample from the patient with a
PT   bladder cancer-associated polynucleotide or antibody.
XX
PS   Claim 10; Page 240-241; 307pp; English.
XX
CC   The present invention describes a method for detecting a bladder cancer-
CC   associated transcript in a cell from a patient. The method comprises
CC   contacting a biological sample from the patient with a polynucleotide
CC   that selectively hybridises to a sequence that is 80 % identical to a
CC   table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC   encode the human bladder cancer-associated proteins given in ABR48146 to
CC   ABR48242). Bladder cancer-associated sequences from the present invention
CC   have cytostatic activities, and can be used in antisense gene therapy and
CC   in vaccine production. The method can be used for detecting a bladder
CC   cancer-associated transcript in a cell from a patient. The method is
CC   useful in diagnosing or treating bladder cancer and in screening for
CC   compounds that modulate bladder cancer, such as hormones or antibodies.
CC   The nucleic acid molecules from the present invention may be used in
CC   various screening and diagnostic methods, and for gene therapy, vaccine
CC   and/or antisense/inhibition applications
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 6;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||     :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:||||   ||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 32
ADF61840
ID   ADF61840 standard; protein; 403 AA.
XX
AC   ADF61840;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Human serine/threonine kinase 15 protein.
XX
KW   cell cycle arrest; cytostatic; antipsoriatic; antiarteriosclerotic;
KW   vasotropic; antithyroid; melanoma; breast; ovarian; lung;
KW   gastrointestinal; colon cancer; Grave's disease; psoriasis;
KW   atherosclerosis; restenosis; vasoproliferative; human;
KW   serine/threonine kinase 15; STK15; ARK2; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO2003088910-A2.
XX
PD   30-OCT-2003.
XX
PF   15-APR-2003; 2003WO-US011867.
XX
PR   15-APR-2002; 2002US-00123568.
```

```
PR   15-APR-2002; 2002US-00123731.
PR   16-APR-2002; 2002US-0373366P.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Hitoshi Y,  Jenkins Y;
XX
DR   WPI; 2003-865396/80.
DR   N-PSDB; ADF61839.
XX
PT   Identifying a compound that modulates cell cycle arrest, for treating
PT   e.g. cancer, comprises contacting a cell comprising a target polypeptide
PT   and determining the chemical or phenotypic effect of the compound upon
PT   the cell.
XX
PS   Claim 1; SEQ ID NO 24; 176pp; English.
XX
CC   The invention relates to a novel method for identifying a compound that
CC   modulates cell cycle arrest comprising contacting a target polypeptide
CC   within a cell with a compound and determining the chemical or phenotypic
CC   effect of the compound upon the cell. The method of the invention has
CC   cytostatic, antipsoriatic, antiarteriosclerotic, vasotropic and
CC   antithyroid applications and may be useful for identifying a compound
CC   that modulates cell cycle arrest. Such compounds may subsequently be used
CC   for developing therapeutic reagents to treat melanoma, breast, ovarian,
CC   lung, gastrointestinal or colon cancer, as well as other proliferative
CC   diseases such as Grave's disease, psoriasis, atherosclerosis, restenosis
CC   and other vasoproliferative diseases. The current sequence is that of the
CC   human serine/threonine kinase 15 (STK15;ARK2) protein of the invention.
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 7;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||   |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 33
ADN39888
ID   ADN39888 standard; protein; 403 AA.
XX
AC   ADN39888;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C258.
XX
KW   Human; differential expression; cancer; angiogenic disorder;
KW   fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW   inflammatory disease; autoimmune disease;
KW   retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW   detection; diagnosis; prognosis; drug screening; drug targeting;
KW   wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW   vulnerary; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003042661-A2.
XX
PD   22-MAY-2003.
XX
PF   13-NOV-2002; 2002WO-US036810.
XX
PR   13-NOV-2001; 2001US-0350666P.
PR   21-NOV-2001; 2001US-0332464P.
PR   29-NOV-2001; 2001US-0334393P.
PR   03-DEC-2001; 2001US-0335394P.
PR   14-DEC-2001; 2001US-0340376P.
PR   08-JAN-2002; 2002US-0347211P.
PR   10-JAN-2002; 2002US-0347349P.
PR   08-FEB-2002; 2002US-0355250P.
PR   13-FEB-2002; 2002US-0356714P.
PR   20-FEB-2002; 2002US-0359077P.
PR   29-MAR-2002; 2002US-0368809P.
PR   04-APR-2002; 2002US-0370110P.
PR   12-APR-2002; 2002US-0372246P.
PR   05-JUN-2002; 2002US-0386614P.
PR   16-JUL-2002; 2002US-0396839P.
PR   22-JUL-2002; 2002US-0397775P.
PR   22-JUL-2002; 2002US-0397845P.
PR   09-SEP-2002; 2002US-0409450P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D,  Aziz N,  Ginsburg WM,  Gish KC,  Glynne R,  Hevezi PA;
PI   Mack DH,  Murray R,  Watson SR,  Wilson KE,  Zlotnik A;
XX
DR   WPI; 2003-468649/44.
DR   N-PSDB; ADN39671.
XX
PT   Determining the presence or absence of a pathological cell in a patient,
PT   useful for diagnosing, prognosing or treating cancer, comprises detecting
PT   a nucleic acid in a biological sample.
XX
PS   Claim 12; SEQ ID NO C258; 1385pp; English.
XX
CC   The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC   whose expression is upregulated or downregulated in specific cancers or
CC   other diseases such as angiogenic or fibrotic disorders, and to methods
CC   of determining the presence or absence of a pathological cell in a
CC   patient by detecting a nucleic acid at least 80% identical to those of
CC   the invention or by detecting a polypeptide of the invention. The
CC   invention also relates to expression vectors and host cells comprising a
CC   nucleic acid of the invention; antibodies which specifically bind a
CC   polypeptide of the invention; use of such antibodies for drug targeting;
CC   and methods of screening for modulators of activity or expression of the
CC   polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC   antibodies and methods are useful for diagnosing, prognosing and treating
CC   cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC   atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC   neovascularistaion syndromes, scarring and uterine fibroids. They may
CC   also be useful in wound healing and in contraception. The present
CC   sequence represents a polypeptide of the invention.
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 7;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
```

```
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||| ||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 34
ADE15337
ID   ADE15337 standard; protein; 403 AA.
XX
AC   ADE15337;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Truncated Aurora-2 kinase #SEQ ID 1.
XX
KW   Aurora-2 kinase; three dimensional structure; protein co-ordinate data.
XX
OS   Homo sapiens.
XX
PN   WO2003092607-A2.
XX
PD   13-NOV-2003.
XX
PF   01-MAY-2003; 2003WO-US013605.
XX
PR   01-MAY-2002; 2002US-0377510P.
XX
PA   (VERT-) VERTEX PHARM INC.
XX
PI   Cheetham G,  Knegtel R,  Swenson L,  Coll JT,  Renwick S,  Weber P;
XX
DR   WPI; 2004-022617/02.
XX
PT   Crystal useful for screening, designing and evaluating compounds as
PT   agonists or antagonists of Aurora-2 kinases, comprises optionally
PT   phosphorylated Aurora-2 kinase domain.
XX
PS   Example 1; SEQ ID NO 1; 242pp; English.
XX
CC   The invention relates to a crystal comprising an optionally
CC   phosphorylated Aurora-2 kinase domain or its homologue. The protein of
CC   the invention may be used for identifying agonists or antagonists of
CC   Aurora-2 kinases, and for generation of three dimensional structures and
CC   structural co-ordinate information of Aurora family proteins. The
CC   crystalline structure facilitates the designing of selective inhibitors
CC   of Aurora family kinases (particularly Aurora-2 kinase) so that
CC   undesirable side effects associated with non-selective inhibitors can be
CC   avoided. The structural co-ordinates solve the structure of Aurora-2
CC   proteins that have amino acid substitutions, additions and/or mutations,
CC   and serve as additional tools to determine the most efficient binding
CC   interactions. The current sequence represents the Aurora-2 kinase amino
CC   acid sequence.
XX
SQ   Sequence 403 AA;

  Query Match           58.5%;  Score 1057.5;  DB 8;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::||||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||| ||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 35
ADH59556
ID   ADH59556 standard; protein; 403 AA.
XX
AC   ADH59556;
XX
DT   25-MAR-2004  (first entry)
XX
DE   Monoclonal antibody of the invention.
XX
KW   monoclonal antibody; MAb; aurora-A kinase; Cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2003106500-A1.
XX
PD   24-DEC-2003.
XX
PF   12-JUN-2003; 2003WO-FR001772.
XX
PR   12-JUN-2002; 2002FR-00007212.
XX
PA   (CNRS ) CENT NAT RECH SCI.
PA   (FRSA-) ETAB FR DU SANG-BRETAGNE.
XX
PI   Prigent C,  Martin A;
XX
DR   WPI; 2004-071550/07.
DR   N-PSDB; ADH59555.
XX
PT   New monoclonal antibody specific for aurora-A kinase, useful for
PT   diagnosis, prognosis and treatment of solid tumors, also for drug
PT   screening.
XX
PS   Claim 8; SEQ ID NO 2; 38pp; French.
XX
CC   The present invention relates to monoclonal antibody (MAb) that
CC   recognizes specifically the human or murine aurora-A kinase binds to
CC   membranes that contain, can detect and optionally purify by
CC   immunoprecipitation, stains biological tissues where is secreted and does
CC   not inhibit the enzymatic activity of MAb. Cytostatic. MAb are used for
CC   in vitro diagnosis and prognosis of cancers in humans and animals,
CC   particularly breast, gastric and colorectal cancer; for treatment of
CC   these cancers; and to screen for inhibitors of MAb. The present sequence
```

CC represents the monoclonal antibody of the invention.
XX
SQ Sequence 403 AA;

Query Match 58.5%; Score 1057.5; DB 8; Length 403;
Best Local Similarity 63.2%; Pred. No. 1.3e-104;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||| :: |||:|| :|:  ||   |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
```

RESULT 36
ADK67720
ID ADK67720 standard; protein; 403 AA.
XX
AC ADK67720;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human modifier of AXIN (MAX) polypeptide.
XX
KW Human; modifier of AXIN; MAX; cytostatic; gene therapy; protein kinase;
KW enzyme.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 31
FT /note= "Encoded by TTT"
FT Misc-difference 104
FT /note= "Encoded by TC"
FT Misc-difference 128
FT /note= "Encoded by TG"
FT Misc-difference 235
FT /note= "Encoded by AACT"
FT Misc-difference 240
FT /note= "Encoded by TT"
XX
PN WO2004013308-A2.
XX
PD 12-FEB-2004.
XX
PF 06-AUG-2003; 2003WO-US024560.
XX
PR 06-AUG-2002; 2002US-0401534P.
PR 16-SEP-2002; 2002US-0411153P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Gendreau SB, Dora EG, Lickteig K, Amundsen CD;
XX
DR WPI; 2004-157122/15.
DR N-PSDB; ADK67713.
XX
PT Identifying a candidate AXIN pathway-modulating agent, useful in
PT diagnosing and treating cancer, comprises providing an assay system
PT comprising a MAX polypeptide or nucleic acid.
XX
PS Example 1; SEQ ID NO 11; 74pp; English.
XX
CC The present sequence is that of a human modifier of AXIN (MAX)
CC polypeptide characterised as a serine/threonine kinase. Genetic screens
CC were designed to identify modifiers of the axin pathway in Caenorhabditis
CC elegans, where a reduction of function pry-1 (axin) mutant was used.
CC Genes causing altered phenotypes were isolated, and their human
CC homologues, termed MAX, were identified, including the present sequence.
CC These MAX polynucleotides and polypeptides are attractive targets for the
CC treatment of pathologies associated with a defective AXIN signalling
CC pathway, such as cancer. Modulation of MAX or its binding partner is
CC useful for understanding the association of the AXIN pathway and its
CC members in normal and disease conditions and for developing diagnostic
CC and therapeutic modalities for AXIN related pathologies. MAX-modulating
CC agents that act by inhibiting or enhancing MAX expression, directly or
CC indirectly, e.g. by affecting MAX function can be identified using
CC methods of the invention. MAX modulating agents are useful in diagnosis,
CC therapy and pharmaceutical development. Preferred MAX modulating agents
CC include antisense and phosphorothioate morpholino oligomers.
XX
SQ Sequence 403 AA;

Query Match 58.5%; Score 1057.5; DB 8; Length 403;
Best Local Similarity 63.2%; Pred. No. 1.3e-104;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

```
Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:          | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||||   |||||||| : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||| :: |||:|| :|:  ||   |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395
```

RESULT 37
ADM72205
ID ADM72205 standard; protein; 403 AA.
XX
AC ADM72205;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human TASK104 polypeptide.
XX
KW TASK; tumour-associated kinase; cytostatic; tumour antigen;
KW cell proliferative disorder; cancer; transgenic; human.
XX
OS Homo sapiens.
XX

PN WO2004024064-A2.
XX
PD 25-MAR-2004.
XX
PF 05-SEP-2003; 2003WO-US027894.
XX
PR 11-SEP-2002; 2002US-0410166P.
XX
PA (GETH) GENENTECH INC.
XX
PI Desauvage FJ, Wood WI, Zhang Z;
XX
DR WPI; 2004-282985/26.
DR N-PSDB; ADM72204.
XX
PT New tumor-associated kinase nucleic acids and polypeptides, useful as
PT hybridization probes for isolating full length TASK DNA, for generating
PT transgenic animals, in chromosome identification, or for tissue typing.
XX
PS Claim 12; SEQ ID NO 10; 163pp; English.
XX
CC The invention relates to new isolated tumour-associated kinase (TASK)
CC nucleic acid molecules and encoded polypeptides. Cytostatic. The
CC antibody, oligopeptide or organic molecule that binds to the TASK
CC polypeptide are useful for treating a mammal having a tumour comprising
CC cells expressing the polypeptide. Antagonists of TASK are useful for
CC treating or preventing a cell proliferative disorder (e.g. cancer)
CC associated with increased expression or activity of TASK polypeptide. The
CC TASK polynucleotides and polypeptides may be used as hybridization probes
CC for isolating full length TASK DNA, for generating transgenic animals, in
CC chromosome identification, or for tissue typing. The present sequence
CC represents a human TASK polypeptide.
XX
SQ Sequence 403 AA;

Query Match 58.5%; Score 1057.5; DB 8; Length 403;
Best Local Similarity 63.2%; Pred. No. 1.3e-104;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

Qy 22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
Db 82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy 74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
Db 130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy 134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
Db 190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy 194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
Db 250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy 254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
Db 310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy 314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
Db 370 QRPMLREVLEHPWITANSSK---PSNCQN 395

RESULT 38
ADQ89832
ID ADQ89832 standard; protein; 403 AA.
XX
AC ADQ89832;
XX
DT 21-OCT-2004 (first entry)
XX
DE Antagonist of cell cycle progression polypeptide #131.
XX
KW Cytostatic; cancer; cell division cycle; mitosis; meiosis;
KW cell cycle progression.
XX
OS Homo sapiens.
XX
PN WO2004063362-A2.
XX
PD 29-JUL-2004.
XX
PF 31-DEC-2003; 2003WO-GB005635.
XX
PR 10-JAN-2003; 2003US-0439123P.
PR 06-MAY-2003; 2003US-0468402P.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Glover D, Bell G, Frenz L, Midgley C;
XX
DR WPI; 2004-544089/52.
DR N-PSDB; ADQ89831.
XX
PT New cell cycle progression genes and proteins for modulating cell cycle
PT progression in cells, for preventing, treating or diagnosing cell
PT proliferative diseases (e.g. cancer) or for identifying modulators of
PT mitosis or meiosis.
XX
PS Claim 2; SEQ ID NO 262; 461pp; English.
XX
CC The present invention relates to a polynucleotide for preventing,
CC treating or diagnosing a disease in an individual. The composition or the
CC polypeptide, polynucleotide or RNA precursor, or antibody is useful for
CC diagnosing, preventing or treating diseases (e.g. cell proliferative
CC diseases such as cancer) in an individual. These may also be used for
CC identifying substances capable of binding to or modulating the function
CC of the polypeptide, capable of affecting the function of the
CC corresponding gene, or capable of inhibiting the cell division cycle or
CC cell cycle progression, preferably mitosis and/or meiosis. The present
CC sequence represents an antagonist of cell cycle progression protein
CC sequence.
XX
SQ Sequence 403 AA;

Query Match 58.5%; Score 1057.5; DB 8; Length 403;
Best Local Similarity 63.2%; Pred. No. 1.3e-104;
Matches 208; Conservative 41; Mismatches 57; Indels 23; Gaps 5;

Qy 22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
Db 82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy 74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
Db 130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy 134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
Db 190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy 194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
Db 250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy 254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
Db 310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy 314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
Db 370 QRPMLREVLEHPWITANSSK---PSNCQN 395

RESULT 39

```
ADR05174
ID   ADR05174 standard; protein; 403 AA.
XX
AC   ADR05174;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Human GTPase regulator-associated w focal adhesion kinase pp125 protein.
XX
KW   apoptosis; cytostatic; antiinflammatory; antiasthmatic; respiratory;
KW   antirheumatic; antiarthritic; gynaecological; cardiant; vasotropic;
KW   antipsoriatic; antiulcer; gastrointestinal; immunosuppressive;
KW   neuroprotective; cancer; autoimmune; neurodegenerative; inflammatory;
KW   asthma; chronic obstructive pulmonary disease; cystic fibrosis;
KW   rheumatoid arthritis; acute respiratory distress syndrome; preeclampsia;
KW   myocardial ischaemia; reperfusion injury; psoriasis; bronchiolitis;
KW   Crohn's disease; ulcerative colitis; inflammatory bowel disease; human;
KW   enzyme; GTPase regulator-associated with focal adhesion kinase pp125;
KW   GRAF.
XX
OS   Homo sapiens.
XX
PN   WO2004065959-A2.
XX
PD   05-AUG-2004.
XX
PF   23-JAN-2004; 2004WO-GB000271.
XX
PR   23-JAN-2003; 2003GB-00001566.
PR   25-MAR-2003; 2003US-0457533P.
XX
PA   (EIRX-) EIRX THERAPEUTICS LTD.
XX
PI   Seery L,  Hayes I,  Murphy F;
XX
DR   WPI; 2004-593556/57.
DR   N-PSDB; ADR05175.
XX
PT   Identifying a modulator of apoptosis-associated polypeptide function,
PT   useful for treating e.g., cancer, comprises incubating a sample
PT   containing an apoptosis-associated polypeptide and a candidate agent to
PT   permit binding.
XX
PS   Claim 1; Page; 230pp; English.
XX
CC   The invention relates to a novel method for identifying an agent that
CC   modulates the function of an apoptosis-associated polypeptide,
CC   particularly a kinase or GCPR (G-protein-coupled receptor). The method
CC   comprises providing a sample containing an apoptosis-associated
CC   polypeptide and a candidate agent and incubating under conditions to
CC   permit binding of the candidate agent to the polypeptide, measuring the
CC   binding and comparing it with the binding of the polypeptide to a control
CC   agent known not to bind to the polypeptide. The method of the invention
CC   has cytostatic, antiinflammatory, antiasthmatic, respiratory,
CC   antirheumatic, antiarthritic, gynaecological, cardiant, vasotropic,
CC   antipsoriatic, antiulcer, gastrointestinal, immunosuppressive and
CC   neuroprotective applications. The method and molecules may be useful for
CC   treating a disease or condition characterised by abnormal apoptosis in
CC   mammalian tissue, particularly cancer, such as small cell lung cancer,
CC   cancer of the kidney, uterus, prostate, bladder, ovary, colon and breast,
CC   leukaemias, sarcomas and myelomas. Furthermore, autoimmune,
CC   neurodegenerative and inflammatory conditions may be treated, including
CC   asthma, chronic obstructive pulmonary disease, cystic fibrosis,
CC   rheumatoid arthritis, acute respiratory distress syndrome, preeclampsia,
CC   myocardial ischaemia, reperfusion injury, psoriasis, bronchiolitis,
CC   Crohn's disease, ulcerative colitis and inflammatory bowel disease. The
CC   current sequence is that of a human apoptosis-associated protein of the
CC   invention which was used during siRNA (small interfering RNA)-mediated
CC   gene silencing.
XX
SQ   Sequence 403 AA;

  Query Match             58.5%;  Score 1057.5;  DB 8;  Length 403;
  Best Local Similarity   63.2%;  Pred. No. 1.3e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |        |:|: |||      :| :   |  ||  |:         | :
Db         82 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 129

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        130 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 189

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:||||||| | :|:|||   |||||||  : |||:|| |||
Db        190 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 249

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:|||||||||||||||:||
Db        250 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 309

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        310 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 369

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        370 QRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 40
ABP41958
ID   ABP41958 standard; protein; 420 AA.
XX
AC   ABP41958;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HPCOK03, SEQ ID NO:3090.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive; chromosome 20q13.2-13.3.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US018569.
XX
PR   07-JUN-2000; 2000US-0209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   N-PSDB; ABQ55035.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT   cancer), immune disorders, cardiovascular disorders and neurological
PT   diseases.
XX
PS   Claim 11; SEQ ID NO 3090; 2922pp; English.
XX
```

```
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
CC   disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC   blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC   respiratory disorders, neurological disorders, gastrointestinal disorders
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds which
CC   modulate ovarian antigen expression or activity. The polynucleotides may
CC   further be used for gene therapy, chromosome mapping, in the
CC   identification of individuals and in forensic analysis, and the
CC   polypeptides may be used as food additives or to prepare antibodies
CC   useful in disease diagnosis, drug targeting and phenotyping. The present
CC   sequence represents a human ovarian antigen of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 420 AA;

  Query Match             58.5%;  Score 1057.5;  DB 5;  Length 420;
  Best Local Similarity   63.2%;  Pred. No. 1.4e-104;
  Matches  208;  Conservative   41;  Mismatches   57;  Indels   23;  Gaps    5;

Qy         22 STLPQRVLR--------KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFT 73
              :::|  | |         |:|: |||      :| :   |  ||  |:         | :
Db         99 TSVPHPVSRPLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK-------RQWA 146

Qy         74 IDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLH 133
              ::|||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||:||
Db        147 LEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLR 206

Qy        134 HPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHG 193
              |||||||| ||:|  |:|||||||| | :|:|||   |||||||| : |||:|| |||
Db        207 HPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHS 266

Qy        194 KKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEK 253
              |:|||||||||||||  |||||||||||||||| || |:||||||||||||||||:||
Db        267 KRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGRMHDEK 326

Qy        254 VDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPS 313
              |||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||
Db        327 VDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPS 386

Qy        314 ERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :|  | :|  |||: ||| :   ||  |:
Db        387 QRPMLREVLEHPWITANSSK---PSNCQN 412


RESULT 41
ABP97470
ID   ABP97470 standard; protein; 319 AA.
XX
AC   ABP97470;
XX
DT   19-SEP-2003  (first entry)
XX
DE   Truncated mutant human Aurora A kinase, [T287] Aurora A (94-402).
XX
KW   Aurora A kinase; human; catalytic domain; chromosome 20q13; 3D structure;
KW   X-ray crystallography; protein co-ordinate data; serine/threonine kinase;
KW   cell cycle regulation; G2/M boundary; oncogenesis;
KW   hyperproliferative disorder; cancer; tumour; breast; colon; colorectal;
KW   Aurora homologue design; rational drug design; drug screening; enzyme;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .10
FT                   /note= "pTB375NBSE vector-encoded sequence, comprising a
FT                   hexahistidine sequence"
FT   Domain          42. .125
FT                   /label= N-terminal_domain
FT                   /note= "Corresponds to residues 125-208 of the native
FT                   human Aurora A kinase"
FT   Active-site     126. .131
FT                   /label= ATP_binding_pocket
FT                   /note= "Corresponds to residues 209-214 of the native
FT                   human Aurora A kinase. Also referred to as a hinge region
FT                   between the N- and C-terminal domains"
FT   Binding-site    127
FT                   /note= "The peptide carbonyl group of this residue binds
FT                   adenine nitrogen atom N6. Corresponds to residue 210 of
FT                   the native human Aurora A kinase"
FT   Binding-site    129
FT                   /note= "The main chain nitrogen of this residue binds
FT                   adenine nitrogen atom N1. Corresponds to residue 212 of
FT                   the native human Aurora A kinase"
FT   Domain          132. .291
FT                   /label= C-terminal_domain
FT                   /note= "Corresponds to residues 215-374 of the native
FT                   human Aurora A kinase"
FT   Misc-difference 204
FT                   /note= "Asp replaces wild-type Thr. Corresponds to
FT                   position 287 of the native human Aurora A kinase"
XX
PN   WO2003031606-A2.
XX
PD   17-APR-2003.
XX
PF   08-OCT-2002; 2002WO-GB004589.
XX
PR   10-OCT-2001; 2001GB-00024299.
XX
PA   (ASTR ) ASTRAZENECA AB.
PA   (ASTR ) ASTRAZENECA UK LTD.
XX
PI   Anderson M,  Keen NJ,  Pannifer ADB,  Pauptit RA,  Rowsell S;
XX
DR   WPI; 2003-403104/38.
XX
PT   Novel crystalline form of polypeptide comprising catalytic domain of
PT   Aurora kinase, whose atomic coordinate data is useful for designing or
PT   selecting Aurora chemical inhibitor, and for designing an Aurora protein.
XX
PS   Example 1; Page 117; 137pp; English.
XX
CC   The invention relates to the three-dimensional structure of human Aurora
CC   A kinase. The invention describes the methods used to prepare and
CC   crystallise truncated mutant forms of human Aurora A kinase comprising
CC   the catalytic domain, which were then subjected to X-ray crystallography
CC   to generate atomic coordinates. The mutant forms of Aurora A used in the
CC   invention all contain a T287D substitution, which produces a mimic of the
CC   activated protein which can be provided as a homogeneous sample (native
CC   Aurora A kinase is activated by phosphorylation of T287). Two crystalline
CC   forms of the Aurora A kinase catalytic domain were produced - one
CC   complexed with the hydrolysable ATP analogue ANP-PNP, and one complexed
CC   with a synthetic inhibitor. The ATP analogue was found to occupy a cleft
CC   between the N-terminal domain (residues 125-208) and the C-terminal
CC   domain (residues 215-374). Aurora A kinase is a serine/threonine kinase
CC   which is involved in cell cycle regulation, with its expression and
CC   activity peaking at the G2/M boundary. It has been implicated in
```

```
CC   oncogenesis - the Aurora A gene maps to chromosome 20q13, a region which
CC   is frequently amplified in human tumours including breast and colon
CC   tumours, and it is overexpressed in over 50% of primary colorectal
CC   tumours. The crystal structure of human Aurora A kinase may be used to
CC   select or design chemical modulators of Aurora kinase, particularly
CC   Aurora kinase inhibitors. These modulators may be used to prevent or
CC   treat undesirable physical and pharmacological consequences of
CC   inappropriate Aurora activity, especially cancers or other
CC   hyperproliferative disorders. The Aurora A kinase 3D structure may also
CC   be used in designing an Aurora protein or Aurora homologue, and in
CC   elucidating the 3D structure of other proteins with structural similarity
CC   to Aurora kinases. The present sequence represents a truncated T287D
CC   mutant Aurora A kinase sequence which was prepared and crystallised in an
CC   example from the invention. This sequence comprises the Aurora A kinase
CC   domain and also contains a vector-encoded sequence at the N-terminus
XX
SQ   Sequence 319 AA;

  Query Match             58.1%;  Score 1050.5;  DB 6;  Length 319;
  Best Local Similarity   65.4%;  Pred. No. 5.3e-104;
  Matches  204;  Conservative   38;  Mismatches   55;  Indels   15;  Gaps    4;

Qy         31 KEPVTPSALVLMSRSNVQPTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGN 90
              |:|: |||      :| :   |  ||  |:         | : ::|||||||||||||||
Db         15 KQPL-PSA----PENNPEEELASKQKNEESKK-------RQWALEDFEIGRPLGKGKFGN 62

Qy         91 VYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRI 150
              |||||||:| ||:|||||||:|:|| |||||||||:|||:|| |||||||| ||:|  |:
Db         63 VYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRV 122

Qy        151 YLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGL 210
              |||||||| | :|:||||   |||||||| : |||:|| ||| |:|||||||||||||
Db        123 YLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALSYCHSKRVIHRDIKPENLLLGS 182

Qy        211 KGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGN 270
              |||||||||||||||| ||  :|||||||||||||||||:|||||| :|||||| |||
Db        183 AGELKIADFGWSVHAPSSRRTDLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGK 242

Qy        271 PPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRAN 330
              ||||: :: |||:|| :|:  ||  |  ||:||||:||:||||:|  | :|  |||: ||
Db        243 PPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITAN 302

Qy        331 SRRVLPPSALQS 342
              | :  ||  |:
Db        303 SSK---PSNCQN 311


RESULT 42
ADJ31727
ID   ADJ31727 standard; protein; 403 AA.
XX
AC   ADJ31727;
XX
DT   22-APR-2004  (first entry)
XX
DE   Human mitotic kinase hARK.
XX
KW   Human; enzyme; SAK; Snk Akin Kinase; serine/threonine protein kinase;
KW   mitotic kinase; cellular proliferation; chemosensitivity; cancer;
KW   proliferative disorder; cytostatic; hARK.
XX
OS   Homo sapiens.
XX
PN   US2003027756-A1.
XX
PD   06-FEB-2003.
XX
PF   21-DEC-2001; 2001US-00026021.
XX
PR   01-AUG-2001; 2001US-0309632P.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Hitoshi Y,  Demo S,  Jenkins Y;
XX
DR   WPI; 2003-479546/45.
XX
PT   Modulating cellular proliferation in subject, by administering
PT   serine/threonine protein kinase involved in modulation of cellular
PT   proliferation and cell cycle regulation, or a nucleic acid encoding the
PT   polypeptide.
XX
PS   Example 1; Fig 2; 41pp; English.
XX
CC   The invention relates to modulating (M1) cellular proliferation in a
CC   subject, involves administering to the subject a serine/threonine protein
CC   kinase involved in modulation of cellular proliferation and cell cycle
CC   regulation (SAK polypeptide, Snk Akin Kinase, a mitotic kinase) encoded
CC   by a nucleic acid that hybridises under stringent conditions to a nucleic
CC   acid encoding a polypeptide appearing as ADJ31723. Also included are
CC   identifying (M2) a compound capable of interfering with binding of SAK or
CC   its fragment (by combining SAK or its fragment with Chk2 polypeptide and
CC   the compound, where SAK or its fragment has kinase activity, and
CC   determining the binding of SAK or its fragment to Chk2), identifying (M3)
CC   a compound that modulates cellular proliferation (by contacting the
CC   compound with SAK, and determining the functional effect of the compound
CC   on SAK), identifying (M4) a compound that modulates cellular
CC   proliferation or chemosensitivity, (by contacting the compound with SAK
CC   or its fragment, and determining the physical effect of the compound on
CC   SAK), determining the chemical or phenotypic effect of the compound upon
CC   a cell comprising SAK or its fragment and modulating (M5) cellular
CC   proliferation in a subject (by administering to the subject a compound
CC   identified by M3). M1 or M5 is useful for modulating cellular
CC   proliferation in a subject, preferably a human having cancer. M2 is
CC   useful for identifying a compound capable of interfering with binding of
CC   SAK or its fragment. M3 is useful for identifying a compound that
CC   modulates cellular proliferation. M4 is useful for identifying a compound
CC   that modulates cellular proliferation or chemosensitivity, where the
CC   compound is an antibody, antisense molecule, small organic molecule, or
CC   circular peptide. M1 or M5 or the compound identified by M3 or M4 is
CC   useful for treating cancer and other proliferative disorders. The present
CC   sequence represents a human mitotic kinase similar to SAK.
XX
SQ   Sequence 403 AA;

  Query Match             58.1%;  Score 1050.5;  DB 7;  Length 403;
  Best Local Similarity   61.7%;  Pred. No. 7.6e-104;
  Matches  209;  Conservative   40;  Mismatches   67;  Indels   23;  Gaps    5;

Qy          4 KENSYPWPYGRQTAPSGLSTLPQRVLRKEPVTPSALVLMSRSNVQPTAAPGQKVMENSSG 63
              :: | | |  |         |       |:|: |||      :| :   |  ||  |:
Db         80 QQTSVPHPVSR--------PLNNTQKSKQPL-PSA----PENNPEEELASKQKNEESKK- 125

Qy         64 TPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLR 123
                 | : ::|||||||||||||||||||||:|  |:|||||||:|:|| |||||||
Db        126 ------RQWALEDFEIGRPLGKGKFGNVYLAREKQSKGILALKVLFKAQLEKAGVEHQLR 179

Qy        124 REIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEE 183
              ||:|||:|| |||||||| ||:|  |:|||||||| | :|:||||   |||||||| : |
Db        180 REVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITE 239

Qy        184 LADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPE 243
              ||:|| ||| |:|||||||||||||  |||||||||||||| || |:||||||||||
Db        240 LANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPE 299

Qy        244 MIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDL 303
              |||||||:|||||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||
Db        300 MIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDL 359

Qy        304 ISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              ||:||:||||:|  | :|  |||: ||| :   ||  |:
Db        360 ISRLLKHNPSQRPMLREVLEHPWITANSSK---PSNCQN 395


RESULT 43
ABR61579
```

```
ID   ABR61579 standard; protein; 402 AA.
XX
AC   ABR61579;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Human HsAIRK-1 (Aurora B) protein.
XX
KW   RasGAP; Aurora kinase; cytostatic; anti-tumoural; human; HsAIRK-1;
KW   Aurora B.
XX
OS   Homo sapiens.
XX
PN   WO2003087395-A2.
XX
PD   23-OCT-2003.
XX
PF   15-APR-2003; 2003WO-IB002972.
XX
PR   15-APR-2002; 2002US-0372483P.
XX
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Garbay C,  Gigoux V,  Camonis J,  L'hoste S,  Samson J;
XX
DR   WPI; 2003-845340/78.
DR   N-PSDB; ACF58068.
XX
PT   Identifying anti-tumoral compounds, comprises determining the capacity of
PT   a compound to inhibit interaction between RasGAP and Drosophila
PT   melanogaster (Dm) aurora kinase, or between RasGAP, Dm aurora kinase and
PT   aurora binding protein.
XX
PS   Disclosure; Page 59-60; 0pp; English.
XX
CC   The invention relates to identifying a biologically active compound with
CC   anti-tumoural properties, where the compound is studied for its capacity
CC   to inhibit the interaction between: (a) RasGAP and the Drosophila
CC   melanogaster Aurora kinase or an orthologue of the kinase, or its
CC   fragment able to interact with RasGAP; or (b) RasGAP, an Aurora-binding
CC   protein and the D. melanogaster Aurora kinase or an orthologue of the
CC   kinase, or its fragment. The method is useful for identifying anti-
CC   tumoural agents. The present sequence represents a human HsAIRK-1 (Aurora
CC   B) protein, an orthologue of the D. melanogaster Aurora kinase
XX
SQ   Sequence 402 AA;

  Query Match             57.8%;  Score 1045.5;  DB 7;  Length 402;
  Best Local Similarity   54.8%;  Pred. No. 2.6e-103;
  Matches  216;  Conservative   44;  Mismatches   75;  Indels   59;  Gaps    8;

Qy          4 KENSYPWPYGRQTAPSG------------LSTLP------QRVL------RKEPVTPSA 38
              |||     |  : ||| |              : ||      ||||      :: |:
Db          5 KENCISGPV-KATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQAQK 63

Qy         39 LV-------------LMSRSNVQPTAAPGQKVMENSSGTPD----ILTRHF--------- 72
              ||             | : |    | : |     ::    |     || |::
Db         64 LVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSHLKIILRRNWHQNRKMKNQ 123

Qy         73 ----TIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEI 128
                  ::||||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:||
Db        124 KEAVALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEI 183

Qy        129 QAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADAL 188
              |:|| |||||||| ||:|  |:|||||||| | :|:||||   ||||||| :   :|:||
Db        184 QSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTANLYNRIANAL 243

Qy        189 MYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGR 248
                ||| |:|||||||||||||  |||||||||||||||| || |:|||||||||||||
Db        244 SYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGR 303

Qy        249 MHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLL 308
              ||:|||||| :|||||| ||| ||||: :: |||:|| :|:  ||  |  ||:||||:||
Db        304 MHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLL 363

Qy        309 RHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
                :||||:|  | :|  |||: ||| :   ||  |:
Db        364 KHNPSQRPMLREVLEHPWITANSSK---PSNCQN 394


RESULT 44
ADD89969
ID   ADD89969 standard; protein; 402 AA.
XX
AC   ADD89969;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human cancer-associated protein kinase STK6.
XX
KW   STK6; protein kinase; enzyme; cancer; cytostatic; immunosuppressive;
KW   antidiabetic; antirheumatic; antiarthritic; antipsoriatic;
KW   antiangiogenic; antiarteriosclerotic; antiinflammatory; vulnerary;
KW   gynaecological; neuroprotective; gene therapy; human.
XX
OS   Homo sapiens.
XX
PN   WO2003083096-A2.
XX
PD   09-OCT-2003.
XX
PF   21-MAR-2003; 2003WO-CA000409.
XX
PR   28-MAR-2002; 2002US-0368853P.
XX
PA   (KINE-) KINETEK PHARM INC.
XX
PI   Delaney AD;
XX
DR   WPI; 2003-833542/77.
DR   N-PSDB; ADD89968.
XX
PT   New nucleic acids encoding cancer associated protein kinases, useful as
PT   targets for screening pharmaceutical agents that inhibit the growth of
PT   tumor cells, or for diagnosing and treating cancer, inflammation or
PT   autoimmune disease.
XX
PS   Claim 1; Page 73-74; 91pp; English.
XX
CC   The present sequence is the protein sequence of the human cancer-
CC   associated protein kinase, STK6. Serine/threonine kinase 6 (STK6),
CC   otherwise known as AIK, is an aurora/IPL1-like kinase possibly involved
CC   in centrosome function. STK6 gene expression is significantly up-
CC   regulated in cancers of the liver, muscle, placenta, and prostate. STK6
CC   is one of a set of protein kinases that are shown by the invention to be
CC   over-expressed in hyper-proliferative cells. These protein kinases
CC   provide targets for drug screening for agents effective in inhibiting the
CC   growth or metastasis of tumour cells, and for determining other molecular
CC   targets in kinase signal transduction pathways involved in transformation
CC   and growth of tumour cells. A claimed method for inhibiting the growth of
CC   a cancer cell involves down-regulating the activity of the protein kinase
CC   using an antisense sequence or inhibitor of kinase activity, especially
CC   where the cancer cell is a breast, liver, colon, muscle, prostate,
CC   kidney, lung, placental or uterine cancer cell. Detection of over-
CC   expression in cancers provides a useful diagnostic for predicting patient
CC   prognosis and probability of drug effectiveness. Agents that specifically
CC   bind the protein kinases can be used for treatment and visualisation of
CC   tumours in patients. The protein kinase polypeptides and nucleic acids
CC   may also be used for treating hyperproliferative diseases, such as
CC   autoimmune disease, diabetes mellitus, multiple sclerosis, rheumatoid
CC   arthritis, psoriasis, atherosclerosis, inflammation, scarring,
CC   endometriosis and angiogenesis.
XX
SQ   Sequence 402 AA;

  Query Match             57.8%;  Score 1045.5;  DB 7;  Length 402;
```

```
  Best Local Similarity   54.8%;  Pred. No. 2.6e-103;
  Matches  216;  Conservative   44;  Mismatches   75;  Indels   59;  Gaps    8;

Qy          4 KENSYPWPYGRQTAPSG-------------LSTLP------QRVL------RKEPVTPSA 38
              |||     |  : ||| |              : ||       ||||      :: |:
Db          5 KENCISGPV-KATAPVGGPKRVLVTQQFPCQNPLPVNSGQAQRVLCPSNSSQRVPLQAQK 63

Qy         39 LV-------------LMSRSNVQPTAAPGQKVMENSSGTPD----ILTRHF--------- 72
              ||              | : |    | : |      ::     |      || |::
Db         64 LVSSHKPVQNQKQKQLQATSVPHPVSRPLNNTQKSKQPLPSHLKIILRRNWHQNRKMKNQ 123

Qy         73 ----TIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEI 128
                  ::||||||||||||||||||||||:| ||:|||||||:|:|| ||||||||||:||
Db        124 KEAVALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEI 183

Qy        129 QAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADAL 188
              |:|| |||||||| ||:|  |:|||||||| | :|:||||   ||||||| :   :|:||
Db        184 QSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTANLYNRIANAL 243

Qy        189 MYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGR 248
               ||| |:|||||||||||||  ||||||||||||||| || |:|||||||||||||||
Db        244 SYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTTLCGTLDYLPPEMIEGR 303

Qy        249 MHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLL 308
              ||:|||||| :|||||| ||| |||| :: |||:|| :|:  ||   |  ||:||||:||
Db        304 MHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLL 363

Qy        309 RHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
              :||||:|  | :|  |||: ||| :   ||  |:
Db        364 KHNPSQRPMLREVLEHPWITANSSK---PSNCQN 394


RESULT 45
ABP97472
ID   ABP97472 standard; protein; 309 AA.
XX
AC   ABP97472;
XX
DT   19-SEP-2003  (first entry)
XX
DE   Truncated mutant human Aurora A kinase, [T287] Aurora A (113-400).
XX
KW   Aurora A kinase; human; catalytic domain; chromosome 20q13; 3D structure;
KW   X-ray crystallography; protein co-ordinate data; serine/threonine kinase;
KW   cell cycle regulation; G2/M boundary; oncogenesis;
KW   hyperproliferative disorder; cancer; tumour; breast; colon; colorectal;
KW   Aurora homologue design; rational drug design; drug screening; enzyme;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .21
FT                   /note= "pET28a vector-encoded sequence, comprising a
FT                   hexahistidine sequence"
FT   Domain          33. .117
FT                   /label= N-terminal_domain
FT                   /note= "Corresponds to residues 125-208 of the native
FT                   human Aurora A kinase"
FT   Active-site     118. .123
FT                   /label= ATP_binding_pocket
FT                   /note= "Corresponds to residues 209-214 of the native
FT                   human Aurora A kinase. Also referred to as a hinge region
FT                   between the N- and C-terminal domains"
FT   Binding-site    119
FT                   /note= "The peptide carbonyl group of this residue binds
FT                   adenine nitrogen atom N6. Corresponds to residue 210 of
FT                   the native human Aurora A kinase"
FT   Binding-site    121
FT                   /note= "The main chain nitrogen of this residue binds
FT                   adenine nitrogen atom N1. Corresponds to residue 212 of
FT                   the native human Aurora A kinase"
FT   Domain          124. .283
FT                   /label= C-terminal_domain
FT                   /note= "Corresponds to residues 215-374 of the native
FT                   human Aurora A kinase"
FT   Misc-difference 196
FT                   /note= "Asp replaces wild-type Thr. Corresponds to
FT                   position 287 of the native human Aurora A kinase"
XX
PN   WO2003031606-A2.
XX
PD   17-APR-2003.
XX
PF   08-OCT-2002; 2002WO-GB004589.
XX
PR   10-OCT-2001; 2001GB-00024299.
XX
PA   (ASTR ) ASTRAZENECA AB.
PA   (ASTR ) ASTRAZENECA UK LTD.
XX
PI   Anderson M,  Keen NJ,  Pannifer ADB,  Pauptit RA,  Rowsell S;
XX
DR   WPI; 2003-403104/38.
XX
PT   Novel crystalline form of polypeptide comprising catalytic domain of
PT   Aurora kinase, whose atomic coordinate data is useful for designing or
PT   selecting Aurora chemical inhibitor, and for designing an Aurora protein.
XX
PS   Example 1; Page 119; 137pp; English.
XX
CC   The invention relates to the three-dimensional structure of human Aurora
CC   A kinase. The invention describes the methods used to prepare and
CC   crystallise truncated mutant forms of human Aurora A kinase comprising
CC   the catalytic domain, which were then subjected to X-ray crystallography
CC   to generate atomic coordinates. The mutant forms of Aurora A used in the
CC   invention all contain a T287D substitution, which produces a mimic of the
CC   activated protein which can be provided as a homogeneous sample (native
CC   Aurora A kinase is activated by phosphorylation of T287). Two crystalline
CC   forms of the Aurora A kinase catalytic domain were produced - one
CC   complexed with the hydrolysable ATP analogue ANP-PNP, and one complexed
CC   with a synthetic inhibitor. The ATP analogue was found to occupy a cleft
CC   between the N-terminal domain (residues 125-208) and the C-terminal
CC   domain (residues 215-374). Aurora A kinase is a serine/threonine kinase
CC   which is involved in cell cycle regulation, with its expression and
CC   activity peaking at the G2/M boundary. It has been implicated in
CC   oncogenesis - the Aurora A gene maps to chromosome 20q13, a region which
CC   is frequently amplified in human tumours including breast and colon
CC   tumours, and it is overexpressed in over 50% of primary colorectal
CC   tumours. The crystal structure of human Aurora A kinase may be used to
CC   select or design chemical modulators of Aurora kinase, particularly
CC   Aurora kinase inhibitors. These modulators may be used to prevent or
CC   treat undesirable physical and pharmacological consequences of
CC   inappropriate Aurora activity, especially cancers or other
CC   hyperproliferative disorders. The Aurora A kinase 3D structure may also
CC   be used in designing an Aurora protein or Aurora homologue, and in
CC   elucidating the 3D structure of other proteins with structural similarity
CC   to Aurora kinases. The present sequence represents a truncated T287D
CC   mutant Aurora A kinase sequence which was prepared and crystallised in an
CC   example from the invention. This sequence comprises the Aurora A kinase
CC   domain and also contains a vector-encoded sequence at the N-terminus
XX
SQ   Sequence 309 AA;

  Query Match             57.7%;  Score 1043.5;  DB 6;  Length 309;
  Best Local Similarity   67.5%;  Pred. No. 2.9e-103;
  Matches  195;  Conservative   37;  Mismatches   54;  Indels    3;  Gaps    1;

Qy         54 GQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQI 113
              |  :: :     :   | : ::|||||||||||||||||||||||:| ||:|||||||:|:
Db         18 GSHMLASKQKNEESKKRQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQL 77

Qy        114 EKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFD 173
              || |||||||||:|||:|| |||||||| ||:|  |:|||||||| | :|:||||   ||
Db         78 EKAGVEHQLRREVEIQSHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFD 137
```

```
Qy        174 EQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTM 233
              ||||||  : |||:|| ||| |:|||||||||||||||  ||||||||||||||| ||  :
Db        138 EQRTATYITELANALSYCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTDL 197

Qy        234 CGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFP 293
              ||||||||||||||||:|||||| :|||||| ||| ||||: :: |||:|| :|:  ||
Db        198 CGTLDYLPPEMIEGRMHDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFP 257

Qy        294 ASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
                |  ||:||||:||:||||:|  | :|  |||: ||| :   ||  |:
Db        258 DFVTEGARDLISRLLKHNPSQRPMLREVLEHPWITANSSK---PSNCQN 303
```

RESULT 46
ABP97473
ID ABP97473 standard; protein; 300 AA.
XX
AC ABP97473;
XX
DT 19-SEP-2003 (first entry)
XX
DE Truncated mutant human Aurora A kinase, [T287] Aurora A (122-400).
XX
KW Aurora A kinase; human; catalytic domain; chromosome 20q13; 3D structure;
KW X-ray crystallography; protein co-ordinate data; serine/threonine kinase;
KW cell cycle regulation; G2/M boundary; oncogenesis;
KW hyperproliferative disorder; cancer; tumour; breast; colon; colorectal;
KW Aurora homologue design; rational drug design; drug screening; enzyme;
KW mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	Region	1. .21
FT		/note= "pET28a vector-encoded sequence, comprising a hexahistidine sequence"
FT	Domain	24. .108
FT		/label= N-terminal_domain
FT		/note= "Corresponds to residues 125-208 of the native human Aurora A kinase"
FT	Active-site	109. .114
FT		/label= ATP_binding_pocket
FT		/note= "Corresponds to residues 209-214 of the native human Aurora A kinase. Also referred to as a hinge region between the N- and C-terminal domains"
FT	Binding-site	110
FT		/note= "The peptide carbonyl group of this residue binds adenine nitrogen atom N6. Corresponds to residue 210 of the native human Aurora A kinase"
FT	Binding-site	112
FT		/note= "The main chain nitrogen of this residue binds adenine nitrogen atom N1. Corresponds to residue 212 of the native human Aurora A kinase"
FT	Domain	115. .274
FT		/label= C-terminal_domain
FT		/note= "Corresponds to residues 215-374 of the native human Aurora A kinase"
FT	Misc-difference	187
FT		/note= "Asp replaces wild-type Thr. Corresponds to position 287 of the native human Aurora A kinase"

XX
PN WO2003031606-A2.
XX
PD 17-APR-2003.
XX
PF 08-OCT-2002; 2002WO-GB004589.
XX
PR 10-OCT-2001; 2001GB-00024299.
XX
PA (ASTR) ASTRAZENECA AB.
PA (ASTR) ASTRAZENECA UK LTD.
XX
PI Anderson M, Keen NJ, Pannifer ADB, Pauptit RA, Rowsell S;
XX
DR WPI; 2003-403104/38.
XX
PT Novel crystalline form of polypeptide comprising catalytic domain of
PT Aurora kinase, whose atomic coordinate data is useful for designing or
PT selecting Aurora chemical inhibitor, and for designing an Aurora protein.
XX
PS Example 1; Page 119; 137pp; English.
XX
CC The invention relates to the three-dimensional structure of human Aurora
CC A kinase. The invention describes the methods used to prepare and
CC crystallise truncated mutant forms of human Aurora A kinase comprising
CC the catalytic domain, which were then subjected to X-ray crystallography
CC to generate atomic coordinates. The mutant forms of Aurora A used in the
CC invention all contain a T287D substitution, which produces a mimic of the
CC activated protein which can be provided as a homogeneous sample (native
CC Aurora A kinase is activated by phosphorylation of T287). Two crystalline
CC forms of the Aurora A kinase catalytic domain were produced - one
CC complexed with the hydrolysable ATP analogue ANP-PNP, and one complexed
CC with a synthetic inhibitor. The ATP analogue was found to occupy a cleft
CC between the N-terminal domain (residues 125-208) and the C-terminal
CC domain (residues 215-374). Aurora A kinase is a serine/threonine kinase
CC which is involved in cell cycle regulation, with its expression and
CC activity peaking at the G2/M boundary. It has been implicated in
CC oncogenesis - the Aurora A gene maps to chromosome 20q13, a region which
CC is frequently amplified in human tumours including breast and colon
CC tumours, and it is overexpressed in over 50% of primary colorectal
CC tumours. The crystal structure of human Aurora A kinase may be used to
CC select or design chemical modulators of Aurora kinase, particularly
CC Aurora kinase inhibitors. These modulators may be used to prevent or
CC treat undesirable physical and pharmacological consequences of
CC inappropriate Aurora activity, especially cancers or other
CC hyperproliferative disorders. The Aurora A kinase 3D structure may also
CC be used in designing an Aurora protein or Aurora homologue, and in
CC elucidating the 3D structure of other proteins with structural similarity
CC to Aurora kinases. The present sequence represents a truncated T287D
CC mutant Aurora A kinase sequence which was prepared and crystallised in an
CC example from the invention. This sequence comprises the Aurora A kinase
CC domain and also contains a vector-encoded sequence at the N-terminus
XX
SQ Sequence 300 AA;

Query Match 57.6%; Score 1041.5; DB 6; Length 300;
Best Local Similarity 71.1%; Pred. No. 4.5e-103;
Matches 194; Conservative 33; Mismatches 43; Indels 3; Gaps 1;

```
Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              | : ::|||||||||||||||||||||:| ||:|||||||:|:|| |||||||||:|||
Db         25 RQWALEDFEIGRPLGKGKFGNVYLAREKQSKFILALKVLFKAQLEKAGVEHQLRREVEIQ 84

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
                 :|| |||||||| ||:|  |:|||||||| | :|:|||||   |||||||| : |||:||
Db         85 SHLRHPNILRLYGYFHDATRVYLILEYAPLGTVYRELQKLSKFDEQRTATYITELANALS 144

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
                 ||| |:|||||||||||||  |||||||||||||||| ||  :|||||||||||||||
Db        145 YCHSKRVIHRDIKPENLLLGSAGELKIADFGWSVHAPSSRRTDLCGTLDYLPPEMIEGRM 204

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPASVPTGAQDLISKLLR 309
              |:|||||| :|||||| ||| |||| :: |||:|| :|:  ||  |  ||:||||:||:
Db        205 HDEKVDLWSLGVLCYEFLVGKPPFEANTYQETYKRISRVEFTFPDFVTEGARDLISRLLK 264

Qy        310 HNPSERLPLAQVSAHPWVRANSRRVLPPSALQS 342
                 ||||:|  | :|  |||: ||| :   ||  |:
Db        265 HNPSQRPMLREVLEHPWITANSSK---PSNCQN 294
```

RESULT 47
AAO18740
ID AAO18740 standard; protein; 403 AA.
XX

```
AC   AAO18740;
XX
DT   24-OCT-2002  (first entry)
XX
DE   Human NOV4 protein.
XX
KW   Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
KW   storage disorder; muscle disorder; neurodegenerative disorder; nootropic;
KW   developmental defect; neuroprotective; antiparkinsonian; hypotensive;
KW   hypertensive; haemostatic; cardiant; antianginal; dermatological;
KW   immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
KW   antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;
KW   vulnerary; anorectic; antidiabetic; immunomodulator; antipsoriatic;
KW   nephrotropic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
KW   antiinfertility; antimanic; antidepressant; metabolic; cytostatic;
KW   tranquilizer; analgesic.
XX
OS   Homo sapiens.
XX
PN   WO200257450-A2.
XX
PD   25-JUL-2002.
XX
PF   29-NOV-2001; 2001WO-US048922.
XX
PR   29-NOV-2000; 2000US-0253834P.
PR   30-NOV-2000; 2000US-0250926P.
PR   25-JAN-2001; 2001US-0264180P.
PR   20-AUG-2001; 2001US-0313656P.
PR   05-OCT-2001; 2001US-0327456P.
PR   28-NOV-2001; 2001US-00327456.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Edinger S,  Macdougall JR,  Millet I,  Ellerman K,  Stone DJ;
PI   Gerlach V,  Grosse WM,  Alsobrook JP,  Lepley DM,  Rieger D,  Burgess CE;
PI   Casman SJ,  Spytek KA,  Boldog FL,  Li L,  Padigaru M,  Mishra V;
PI   Patturajan M,  Shenoy S,  Rastelli L,  Tchernev VT,  Vernet CAM;
PI   Zerhusen BD,  Malyankar UM,  Guo X,  Miller CE,  Gangolli EA;
XX
DR   WPI; 2002-590741/63.
DR   N-PSDB; ABT06285.
XX
PT   Novel isolated polypeptide, designated NOVX, useful for treating or
PT   preventing in NOVX-associated disorders e.g. cardiomyopathy,
PT   atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.
XX
PS   Claim 1; Page 62; 353pp; English.
XX
CC   The present invention provides the protein and coding sequences of
CC   several novel human proteins, designated NOVX. These can be used in the
CC   treatment of, amongst others, cancers, autoimmune diseases, infections,
CC   inflammatory diseases, storage disorders, muscle disorders,
CC   neurodegenerative diseases and developmental defects. The present
CC   sequence is a protein of the invention
XX
SQ   Sequence 403 AA;

  Query Match             53.4%;  Score 966;  DB 5;  Length 403;
  Best Local Similarity   61.8%;  Pred. No. 9.9e-95;
  Matches  186;  Conservative   41;  Mismatches   68;  Indels    6;  Gaps    2;

Qy         45 SNVQPTAAP---GQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHF 101
              |   | :||    :: : :     :   | : ::| ||||| |||||||||||||:| |
Db         98 SKQSPLSAPENNPEEELASKQKNEESKKRQWALEDLEIGRPPGKGKFGNVYLAREKQSKF 157

Qy        102 IVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGE 161
              |:||:||||:|:|| |||||||||:|||:|| ||||:||| ||:|  |:|||||| |
Db        158 ILALRVLFKAQLEKAGVEHQLRREVEIQSHLQHPNIIRLYGYFHDATRVYLILEYTPLET 217

Qy        162 LYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGW 221
              :  ||||   |||||||| : ||| || ||| | |||||||||||||  |||:||:|||
Db        218 VNTELQKLSKFDEQRTATYITELASALSYCHSKTVIHRDIKPENLLLGSAGELEIANFGW 277

Qy        222 SVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNET 281
              | |||| || |:|||||||||||||||:|||||| :|||| | ||| ||||: :: ||
Db        278 SEHAPSSRRTTLCGTLDYLPPEMIEGRMHDEKVDLWSLGVLCCEFLVGKPPFEANTYQET 337

Qy        282 YRRIVKVDLKFPASVPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQ 341
              |:|| :|:  ||  |  ||:||||:||:| ||:|  | :|  :||: ||| :   ||  |
Db        338 YKRISRVEFTFPDFVTEGARDLISRLLKHVPSQRPMLREVLEYPWITANSSK---PSNCQ 394

Qy        342 S 342
              :
Db        395 N 395


RESULT 48
ADJ49242
ID   ADJ49242 standard; protein; 294 AA.
XX
AC   ADJ49242;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Oil-associated gene related protein #742.
XX
KW   oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS   Unidentified.
XX
PN   US2004025202-A1.
XX
PD   05-FEB-2004.
XX
PF   14-MAR-2003; 2003US-00389566.
XX
PR   15-MAR-2002; 2002US-0365301P.
PR   26-JUN-2002; 2002US-0391786P.
PR   26-JUN-2002; 2002US-0392018P.
XX
PA   (LAUR/) LAURIE C C.
PA   (RAVA/) RAVANELLO M.
PA   (SAVA/) SAVAGE T.
PA   (LEDE/) LEDEAUX J R.
PA   (ROGE/) ROGERS J A.
XX
PI   Laurie CC,  Ravanello M,  Savage T,  Ledeaux JR,  Rogers JA;
XX
DR   WPI; 2004-142683/14.
XX
PT   Novel recombinant DNA construct comprising a promoter functional in
PT   plants operably linked to an oil-associated gene for producing transgenic
PT   plant seed.
XX
PS   Example 3; SEQ ID NO 1246; 22pp; English.
XX
CC   The invention relates to a recombinant DNA construct comprising a
CC   promoter functional in plants operably linked to an oil-associated gene.
CC   The construct is useful for transgenic plant seed which has in its genome
CC   the construct, that is functional in the plant to transcribe the oil-
CC   associated gene. The transgenic plant seed grows into a plant having
CC   enhanced seed oil as compared to wild type. The construct is useful for
CC   producing hybrid maize seed. The transgenic plant seed is useful for
CC   producing vegetable oil. The present sequence represents the amino acid
CC   sequence of an oil-associated gene related protein.
XX
SQ   Sequence 294 AA;

  Query Match             52.0%;  Score 940;  DB 8;  Length 294;
  Best Local Similarity   60.3%;  Pred. No. 3.9e-92;
  Matches  179;  Conservative   47;  Mismatches   63;  Indels    8;  Gaps    4;

Qy         49 PTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVL 108
              ||    |:  | |  :     : :|: ||:||:|||:||||:|||||||:|: :||||||
Db          4 PTETQHQE-KEASDASAAAAQKRWTLSDFDIGKPLGRGKFGHVYLAREKRSNHVVALKVL 62
```

```
Qy        109 FKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQK 168
              ||||:::  |||||||:|||:|| |||||||| ||||::|:||||||| |||||||:|||
Db         63 FKSQLQQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQKRVYLILEYAARGELYKDLQK 122

Qy        169 SCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSL 228
                | |:| || :  || ||:||||| |||||||||||:| :|||||||||||||  :
Db        123 CKYFSERRAATYVASLARALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN- 181

Qy        229 RRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKV 288
              ||:|||||||||||||:|   |:  ||:| :|:|||| | | ||||:  |::||||||:|
Db        182 RRRTMCGTLDYLPPEMVESVEHDASVDIWSLGILCYEFLYGVPPFEAMEHSDTYRRIVQV 241

Qy        289 DLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
              |||||    :   |:||||::|    |:|||| ::  |||:  |:     || :  |
Db        242 DLKFPPKPIISASAKDLISQMLVKESSQRLPLHKLLEHPWIVQNA----DPSGIYRV 294
```

RESULT 49
ADJ49602
ID ADJ49602 standard; protein; 294 AA.
XX
AC ADJ49602;
XX
DT 06-MAY-2004 (first entry)
XX
DE Oil-associated gene related protein #1102.
XX
KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS Unidentified.
XX
PN US2004025202-A1.
XX
PD 05-FEB-2004.
XX
PF 14-MAR-2003; 2003US-00389566.
XX
PR 15-MAR-2002; 2002US-0365301P.
PR 26-JUN-2002; 2002US-0391786P.
PR 26-JUN-2002; 2002US-0392018P.
XX
PA (LAUR/) LAURIE C C.
PA (RAVA/) RAVANELLO M.
PA (SAVA/) SAVAGE T.
PA (LEDE/) LEDEAUX J R.
PA (ROGE/) ROGERS J A.
XX
PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
XX
DR WPI; 2004-142683/14.
XX
PT Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.
XX
PS Example 3; SEQ ID NO 1606; 22pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in plants operably linked to an oil-associated gene.
CC The construct is useful for transgenic plant seed which has in its genome
CC the construct, that is functional in the plant to transcribe the oil-
CC associated gene. The transgenic plant seed grows into a plant having
CC enhanced seed oil as compared to wild type. The construct is useful for
CC producing hybrid maize seed. The transgenic plant seed is useful for
CC producing vegetable oil. The present sequence represents the amino acid
CC sequence of an oil-associated gene related protein.
XX
SQ Sequence 294 AA;

Query Match 51.8%; Score 937; DB 8; Length 294;
Best Local Similarity 60.3%; Pred. No. 8.1e-92;
Matches 179; Conservative 46; Mismatches 64; Indels 8; Gaps 4;

```
Qy         49 PTAAPGQKVMENSSGTPDILTRHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVL 108
              ||     |:  | |  :      : :|: ||:||:|||:||||:|||||||:|: :||||||
Db          4 PTETQHQE-KEASDASAAAAQKRWTLSDFDIGKPLGRGKFGHVYLAREKRSNHVVALKVL 62

Qy        109 FKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQK 168
              ||||:::  |||||||:|||:|| |||||||| ||||::|:||||||| |||||||:|||
Db         63 FKSQLQQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQKRVYLILEYAARGELYKDLQK 122

Qy        169 SCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSL 228
                | |:| || :  || ||:||||| |||||||||||:| :|||||||||||||  :
Db        123 CKYFSERRAATYVASLARALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFNX 182

Qy        229 RRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKV 288
              || ||||||||||||:|   |:  ||:| :|:|||| | | ||||:  |::||||||:|
Db        183 RR-TMCGTLDYLPPEMVESVEHDASVDIWSLGILCYEFLYGVPPFEAMEHSDTYRRIVQV 241

Qy        289 DLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRANSRRVLPPSALQSV 343
              |||||    :   |:||||::|    |:|||| ::  |||:  |:     || :  |
Db        242 DLKFPPKPIISASAKDLISQMLVKESSQRLPLHKLLEHPWIVQNA----DPSGIYRV 294
```

RESULT 50
ADJ48641
ID ADJ48641 standard; protein; 295 AA.
XX
AC ADJ48641;
XX
DT 06-MAY-2004 (first entry)
XX
DE Oil-associated gene related protein #141.
XX
KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS Unidentified.
XX
PN US2004025202-A1.
XX
PD 05-FEB-2004.
XX
PF 14-MAR-2003; 2003US-00389566.
XX
PR 15-MAR-2002; 2002US-0365301P.
PR 26-JUN-2002; 2002US-0391786P.
PR 26-JUN-2002; 2002US-0392018P.
XX
PA (LAUR/) LAURIE C C.
PA (RAVA/) RAVANELLO M.
PA (SAVA/) SAVAGE T.
PA (LEDE/) LEDEAUX J R.
PA (ROGE/) ROGERS J A.
XX
PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
XX
DR WPI; 2004-142683/14.
XX
PT Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.
XX
PS Example 3; SEQ ID NO 645; 22pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in plants operably linked to an oil-associated gene.
CC The construct is useful for transgenic plant seed which has in its genome
CC the construct, that is functional in the plant to transcribe the oil-
CC associated gene. The transgenic plant seed grows into a plant having
CC enhanced seed oil as compared to wild type. The construct is useful for
CC producing hybrid maize seed. The transgenic plant seed is useful for
CC producing vegetable oil. The present sequence represents the amino acid
CC sequence of an oil-associated gene related protein.
XX
SQ Sequence 295 AA;

```
  Query Match             51.7%;  Score 934.5;  DB 8;  Length 295;
  Best Local Similarity   65.5%;  Pred. No. 1.5e-91;
  Matches  173;  Conservative   44;  Mismatches   44;  Indels    3;  Gaps    2;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              | :|::||:||:|||:||||:||||||| |: |||||||||||:::  | ||||||:|||
Db         25 RRWTLNDFDIGKPLGRGKFGHVYLAREKTSNHIVALKVLFKSQLQQSQVVHQLRREVEIQ 84

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| ||:|||||  ||||::|:|||||||:|||||||    | |:| || :  || ||:
Db         85 SHLRHPHILRLYGYFYDQKRVYLILEYAPKGELYKELQKCKYFSERRAATYVASLARALI 144

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||||| |||||||||||:| :|||||||||||  : ||:|||||||||||:|
Db        145 YCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 203

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|||||| | | ||||:  |::|||||::||||||    | : |:||||::
Db        204 HDASVDIWSLGVLCYEFLYGVPPFEAKEHSDTYRRIIQVDLKFPPKPIVSSAAKDLISQM 263

Qy        308 LRHNPSERLPLAQVSAHPWVRANS 331
              |  : |:|||| ::  |||:  |:
Db        264 LVKDSSQRLPLHKLLEHPWIVQNA 287


RESULT 51
ADJ48640
ID   ADJ48640 standard; protein; 298 AA.
XX
AC   ADJ48640;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Oil-associated gene related protein #140.
XX
KW   oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS   Unidentified.
XX
PN   US2004025202-A1.
XX
PD   05-FEB-2004.
XX
PF   14-MAR-2003; 2003US-00389566.
XX
PR   15-MAR-2002; 2002US-0365301P.
PR   26-JUN-2002; 2002US-0391786P.
PR   26-JUN-2002; 2002US-0392018P.
XX
PA   (LAUR/) LAURIE C C.
PA   (RAVA/) RAVANELLO M.
PA   (SAVA/) SAVAGE T.
PA   (LEDE/) LEDEAUX J R.
PA   (ROGE/) ROGERS J A.
XX
PI   Laurie CC,  Ravanello M,  Savage T,  Ledeaux JR,  Rogers JA;
XX
DR   WPI; 2004-142683/14.
XX
PT   Novel recombinant DNA construct comprising a promoter functional in
PT   plants operably linked to an oil-associated gene for producing transgenic
PT   plant seed.
XX
PS   Example 3; SEQ ID NO 644; 22pp; English.
XX
CC   The invention relates to a recombinant DNA construct comprising a
CC   promoter functional in plants operably linked to an oil-associated gene.
CC   The construct is useful for transgenic plant seed which has in its genome
CC   the construct, that is functional in the plant to transcribe the oil-
CC   associated gene. The transgenic plant seed grows into a plant having
CC   enhanced seed oil as compared to wild type. The construct is useful for
CC   producing hybrid maize seed. The transgenic plant seed is useful for
CC   producing vegetable oil. The present sequence represents the amino acid
CC   sequence of an oil-associated gene related protein.
XX
SQ   Sequence 298 AA;

  Query Match             51.6%;  Score 933.5;  DB 8;  Length 298;
  Best Local Similarity   65.5%;  Pred. No. 2e-91;
  Matches  173;  Conservative   44;  Mismatches   44;  Indels    3;  Gaps    2;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              | :|::||:||:|||:||||:||||||| |: |||||||||||:::  | ||||||:|||
Db         28 RRWTLNDFDIGKPLGRGKFGHVYLAREKTSNHIVALKVLFKSQLQQSQVVHQLRREVEIQ 87

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| ||:|||||  ||||::|:|||||||:|||||||    | |:| || :  || ||:
Db         88 SHLRHPHILRLYGYFYDQKRVYLILEYAPKGELYKELQKCKYFSERRAATYVASLARALI 147

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||||| |||||||||||:| :|||||||||||  : ||:|||||||||||:|
Db        148 YCHGKHVIHRDIKPENLLIGSQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 206

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|||||| | | ||||:  |::|||||::||||||    | : |:||||::
Db        207 HDASVDIWSLGVLCYEFLYGVPPFEAKEHSDTYRRIIQVDLKFPPKPIVSSAAKDLISQM 266

Qy        308 LRHNPSERLPLAQVSAHPWVRANS 331
              |  : |:|||| ::  |||:  |:
Db        267 LVKDSSQRLPLHKLLEHPWIVQNA 290


RESULT 52
ADJ49236
ID   ADJ49236 standard; protein; 282 AA.
XX
AC   ADJ49236;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Oil-associated gene related protein #736.
XX
KW   oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS   Unidentified.
XX
PN   US2004025202-A1.
XX
PD   05-FEB-2004.
XX
PF   14-MAR-2003; 2003US-00389566.
XX
PR   15-MAR-2002; 2002US-0365301P.
PR   26-JUN-2002; 2002US-0391786P.
PR   26-JUN-2002; 2002US-0392018P.
XX
PA   (LAUR/) LAURIE C C.
PA   (RAVA/) RAVANELLO M.
PA   (SAVA/) SAVAGE T.
PA   (LEDE/) LEDEAUX J R.
PA   (ROGE/) ROGERS J A.
XX
PI   Laurie CC,  Ravanello M,  Savage T,  Ledeaux JR,  Rogers JA;
XX
DR   WPI; 2004-142683/14.
XX
PT   Novel recombinant DNA construct comprising a promoter functional in
PT   plants operably linked to an oil-associated gene for producing transgenic
PT   plant seed.
XX
PS   Example 3; SEQ ID NO 1240; 22pp; English.
XX
CC   The invention relates to a recombinant DNA construct comprising a
CC   promoter functional in plants operably linked to an oil-associated gene.
CC   The construct is useful for transgenic plant seed which has in its genome
```

```
CC   the construct, that is functional in the plant to transcribe the oil-
CC   associated gene. The transgenic plant seed grows into a plant having
CC   enhanced seed oil as compared to wild type. The construct is useful for
CC   producing hybrid maize seed. The transgenic plant seed is useful for
CC   producing vegetable oil. The present sequence represents the amino acid
CC   sequence of an oil-associated gene related protein.
XX
SQ   Sequence 282 AA;

  Query Match             50.9%;  Score 921.5;  DB 8;  Length 282;
  Best Local Similarity   63.7%;  Pred. No. 3.6e-90;
  Matches  174;  Conservative   42;  Mismatches   50;  Indels    7;  Gaps    3;

Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              : :|  ||:||:|||:||||:|||||||:|  |||||||||:|:::  ||||||||:|||
Db         12 KRWTTSDFDIGKPLGRGKFGHVYLAREKRSDHIVALKVLFKAQLQQSQVEHQLRREVEIQ 71

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| |||||||| ||||::|:||||||| |||||||||    | |:| || :  || ||:
Db         72 SHLRHPNILRLYGYFYDQKRVYLILEYAVRGELYKELQKCKYFSERRAATYVASLARALI 131

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              ||||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        132 YCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVESVE 190

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|:|||| | | ||||:  |:|||:|||:||||||    | : |:||||::
Db        191 HDASVDIWSLGILCYEFLYGVPPFEAREHSETYKRIVQVDLKFPPKPIVSSSAKDLISQM 250

Qy        308 LRHNPSERLPLAQVSAHPWVRANSRRVLPPSAL 340
              |    ::|| | ::  |||:  |:     || |
Db        251 LVKESTQRLALHKLLEHPWIVQNA----DPSGL 279


RESULT 53
ADJ48378
ID   ADJ48378 standard; protein; 368 AA.
XX
AC   ADJ48378;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Maize oil-associated gene protein #37.
XX
KW   maize; plant; oil-associated gene; transgenic; enhanced seed oil;
KW   vegetable oil.
XX
OS   Zea mays.
XX
PN   US2004025202-A1.
XX
PD   05-FEB-2004.
XX
PF   14-MAR-2003; 2003US-00389566.
XX
PR   15-MAR-2002; 2002US-0365301P.
PR   26-JUN-2002; 2002US-0391786P.
PR   26-JUN-2002; 2002US-0392018P.
XX
PA   (LAUR/) LAURIE C C.
PA   (RAVA/) RAVANELLO M.
PA   (SAVA/) SAVAGE T.
PA   (LEDE/) LEDEAUX J R.
PA   (ROGE/) ROGERS J A.
XX
PI   Laurie CC,  Ravanello M,  Savage T,  Ledeaux JR,  Rogers JA;
XX
DR   WPI; 2004-142683/14.
XX
PT   Novel recombinant DNA construct comprising a promoter functional in
PT   plants operably linked to an oil-associated gene for producing transgenic
PT   plant seed.
XX
PS   Disclosure; SEQ ID NO 382; 22pp; English.
XX
CC   The invention relates to a recombinant DNA construct comprising a
CC   promoter functional in plants operably linked to an oil-associated gene.
CC   The construct is useful for transgenic plant seed which has in its genome
CC   the construct, that is functional in the plant to transcribe the oil-
CC   associated gene. The transgenic plant seed grows into a plant having
CC   enhanced seed oil as compared to wild type. The construct is useful for
CC   producing hybrid maize seed. The transgenic plant seed is useful for
CC   producing vegetable oil. The present sequence represents the amino acid
CC   sequence of a maize oil-associated gene protein.
XX
SQ   Sequence 368 AA;

  Query Match             50.7%;  Score 918;  DB 8;  Length 368;
  Best Local Similarity   58.8%;  Pred. No. 1.3e-89;
  Matches  177;  Conservative   47;  Mismatches   59;  Indels   18;  Gaps    4;

Qy         48 QPT--AAPGQKVMENSSGTPDIL----------TRHFTIDDFEIGRPLGKGKFGNVY 92
              :||  ::|   :: |  |   :              : : : |||:|:|||:||||:||
Db         57 RPTHASSPSAPLLHNHGGRHRVARQRGQGXXSHANQEKRWVLSDFEVGKPLGRGKFGHVY 116

Qy         93 LAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQAHLHHPNILRLYNYFYDRRRIYL 152
              |||||:|  |||||||||:::  ||||||||:|||:|| |||||||: |:||
Db        117 LAREKRSSEIVALKVLFKSQLKQSQVEHQLRREVEIQSHLRHPNILRLYGYFYDQTRVYL 176

Qy        153 ILEYAPRGELYKELQKSCTFDEQRTATIMEELADALMYCHGKKVIHRDIKPENLLLGLKG 212
              ||||| :|||||| :   | |:|||| :  || ||:| ||| |||||||||||:|::|
Db        177 ILEYAAKGELYKELTRCKHFSERRTATYVASLARALIYLHGKHVIHRDIKPENLLVGVQG 236

Qy        213 ELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRMHNEKVDLWCIGVLCYELLVGNPP 272
              |:||||||||||  : ||:|||||||||||||:|   |:  ||:| :|||||| | | ||
Db        237 EIKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVEKAEHDYHVDIWSLGVLCYEFLYGVPP 295

Qy        273 FESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKLLRHNPSERLPLAQVSAHPWVRAN 330
              ||:  |:|||||||||||||    |   |:||||::|  : ::|||| :|  |||:  |
Db        296 FEAKEHSETYRRIVKVDLKFPLKPFVSPAAKDLISQMLVKSSAQRLPLHKVLEHPWIVQN 355

Qy        331 S 331
              :
Db        356 A 356


RESULT 54
ADJ48815
ID   ADJ48815 standard; protein; 418 AA.
XX
AC   ADJ48815;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Oil-associated gene related protein #315.
XX
KW   oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS   Unidentified.
XX
PN   US2004025202-A1.
XX
PD   05-FEB-2004.
XX
PF   14-MAR-2003; 2003US-00389566.
XX
PR   15-MAR-2002; 2002US-0365301P.
PR   26-JUN-2002; 2002US-0391786P.
PR   26-JUN-2002; 2002US-0392018P.
XX
PA   (LAUR/) LAURIE C C.
PA   (RAVA/) RAVANELLO M.
PA   (SAVA/) SAVAGE T.
PA   (LEDE/) LEDEAUX J R.
PA   (ROGE/) ROGERS J A.
XX
```

PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
XX
DR WPI; 2004-142683/14.
XX
PT Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.
XX
PS Example 3; SEQ ID NO 819; 22pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in plants operably linked to an oil-associated gene.
CC The construct is useful for transgenic plant seed which has in its genome
CC the construct, that is functional in the plant to transcribe the oil-
CC associated gene. The transgenic plant seed grows into a plant having
CC enhanced seed oil as compared to wild type. The construct is useful for
CC producing hybrid maize seed. The transgenic plant seed is useful for
CC producing vegetable oil. The present sequence represents the amino acid
CC sequence of an oil-associated gene related protein.
XX
SQ Sequence 418 AA;

Query Match 50.7%; Score 917.5; DB 8; Length 418;
Best Local Similarity 65.2%; Pred. No. 1.8e-89;
Matches 172; Conservative 41; Mismatches 48; Indels 3; Gaps 2;

```
Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              : : : ||:||:|||:||||:|||||||:|: ||||||||||:::  ||||||||:|||
Db        148 KRWVLSDFDIGKPLGRGKFGHVYLAREKRSNHIVALKVLFKSQLKQSQVEHQLRREVEIQ 207

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
              :|| ||||||||| ||||: |:|||||||| :|||||||:   | |:|:|| :  || ||:
Db        208 SHLRHPNILRLYGYFYDQTRVYLILEYALKGELYKELQRCKHFSERRSATYIASLAHALI 267

Qy        190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
              | ||| |||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        268 YLHGKHVIHRDIKPENLLIGSQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVEKTE 326

Qy        250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
              |:  ||:| :|:|||| | | ||||:  |:|||||||||||||    |   |:||||::
Db        327 HDYHVDIWSLGILCYEFLYGVPPFEAKEHSETYRRIVKVDLKFPLKPFVSPAAKDLISQM 386

Qy        308 LRHNPSERLPLAQVSAHPWVRANS 331
              |  | : |||| ::  |||:  |:
Db        387 LVKNSAHRLPLHKLLEHPWIVQNA 410
```

RESULT 55
ADJ48546
ID ADJ48546 standard; protein; 298 AA.
XX
AC ADJ48546;
XX
DT 06-MAY-2004 (first entry)
XX
DE Oil-associated gene related protein #46.
XX
KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS Unidentified.
XX
PN US2004025202-A1.
XX
PD 05-FEB-2004.
XX
PF 14-MAR-2003; 2003US-00389566.
XX
PR 15-MAR-2002; 2002US-0365301P.
PR 26-JUN-2002; 2002US-0391786P.
PR 26-JUN-2002; 2002US-0392018P.
XX
PA (LAUR/) LAURIE C C.
PA (RAVA/) RAVANELLO M.
PA (SAVA/) SAVAGE T.
PA (LEDE/) LEDEAUX J R.
PA (ROGE/) ROGERS J A.
XX
PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
XX
DR WPI; 2004-142683/14.
XX
PT Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.
XX
PS Example 3; SEQ ID NO 550; 22pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in plants operably linked to an oil-associated gene.
CC The construct is useful for transgenic plant seed which has in its genome
CC the construct, that is functional in the plant to transcribe the oil-
CC associated gene. The transgenic plant seed grows into a plant having
CC enhanced seed oil as compared to wild type. The construct is useful for
CC producing hybrid maize seed. The transgenic plant seed is useful for
CC producing vegetable oil. The present sequence represents the amino acid
CC sequence of an oil-associated gene related protein.
XX
SQ Sequence 298 AA;

Query Match 50.2%; Score 908; DB 8; Length 298;
Best Local Similarity 63.7%; Pred. No. 1.1e-88;
Matches 170; Conservative 45; Mismatches 46; Indels 6; Gaps 3;

```
Qy         70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
              | :|::||:||:|||:||||:|||||||  |: ||||||||||:::  | ||||||:|||
Db         25 RRWTLNDFDIGKPLGRGKFGHVYLAREKTSNHIVALKVLFKSQLQQSQVVHQLRREVEIQ 84

Qy        130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTF---DEQRTATIMEELAD 186
              :|| ||:| ||| ||||::|:|||||||:|||||||||   |   : :  || :  ||
Db         85 SHLRHPHIWRLYGYFYDQKRVYLILEYAPKGELYKELQKCKYFSRXERRAXATYVASLAR 144

Qy        187 ALMYCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIE 246
              ||:||||| |||||||||||||:| :|||||||||||||  : ||:|||||||||||||:|
Db        145 ALIYCHGKHVIHRDIKPENLLIGAQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVE 203

Qy        247 GRMHNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLI 304
                 |:  ||:| :|||||| | | ||||:  |::|||||::||||||    | : |:|||
Db        204 SVEHDASVDIWSLGVLCYEFLYGVPPFEAKEHSDTYRRIIQVDLKFPPKPIVSSAAKDLI 263

Qy        305 SKLLRHNPSERLPLAQVSAHPWVRANS 331
              |::|  : |:|||| ::  |||:  |:
Db        264 SQMLVKDSSQRLPLHKLLEHPWIVQNA 290
```

RESULT 56
ADJ50416
ID ADJ50416 standard; protein; 432 AA.
XX
AC ADJ50416;
XX
DT 06-MAY-2004 (first entry)
XX
DE Oil-associated gene related protein #1916.
XX
KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS Unidentified.
XX
PN US2004025202-A1.
XX
PD 05-FEB-2004.
XX
PF 14-MAR-2003; 2003US-00389566.
XX
PR 15-MAR-2002; 2002US-0365301P.
PR 26-JUN-2002; 2002US-0391786P.

PR 26-JUN-2002; 2002US-0392018P.
XX
PA (LAUR/) LAURIE C C.
PA (RAVA/) RAVANELLO M.
PA (SAVA/) SAVAGE T.
PA (LEDE/) LEDEAUX J R.
PA (ROGE/) ROGERS J A.
XX
PI Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
XX
DR WPI; 2004-142683/14.
XX
PT Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.
XX
PS Example 3; SEQ ID NO 2420; 22pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in plants operably linked to an oil-associated gene.
CC The construct is useful for transgenic plant seed which has in its genome
CC the construct, that is functional in the plant to transcribe the oil-
CC associated gene. The transgenic plant seed grows into a plant having
CC enhanced seed oil as compared to wild type. The construct is useful for
CC producing hybrid maize seed. The transgenic plant seed is useful for
CC producing vegetable oil. The present sequence represents the amino acid
CC sequence of an oil-associated gene related protein.
XX
SQ Sequence 432 AA;

Query Match 50.1%; Score 906; DB 8; Length 432;
Best Local Similarity 65.2%; Pred. No. 3.3e-88;
Matches 172; Conservative 40; Mismatches 48; Indels 4; Gaps 3;

```
Qy          70 RHFTIDDFEIGRPLGKGKFGNVYLAREKKSHFIVALKVLFKSQIEKEGVEHQLRREIEIQ 129
               : : : ||:||:|||:||||:||||||||:|: ||||||||||:::  ||||||||:|||
Db         163 KRWVLSDFDIGKPLGRGKFGHVYLAREKRSNHIVALKVLFKSQLKQSQVEHQLRREVEIQ 222

Qy         130 AHLHHPNILRLYNYFYDRRRIYLILEYAPRGELYKELQKSCTFDEQRTATIMEELADALM 189
               :|| |||||||| |||| |:|||||| :|||||||:   | |:|:|| :  || ||:
Db         223 SHLRHPNILRLYGYFYD-TRVYLILEYALKGELYKELQRCKHFSERRSATYIASLAHALI 281

Qy         190 YCHGKKVIHRDIKPENLLLGLKGELKIADFGWSVHAPSLRRKTMCGTLDYLPPEMIEGRM 249
               | ||| |||||||||||:| :||||||||||||  : ||:|||||||||||||:|
Db         282 YLHGKHVIHRDIKPENLLIGSQGELKIADFGWSVHTFN-RRRTMCGTLDYLPPEMVEKTE 340

Qy         250 HNEKVDLWCIGVLCYELLVGNPPFESASHNETYRRIVKVDLKFPAS--VPTGAQDLISKL 307
               |:  ||:| :|:|||| | | ||||:  |:||||||||||||||    |   |:||||::
Db         341 HDYHVDIWSLGILCYEFLYGVPPFEAKEHSETYRRIVKVDLKFPLKPFVSPAAKDLISQM 400

Qy         308 LRHNPSERLPLAQVSAHPWVRANS 331
               |  | : |||| ::  |||:  |:
Db         401 LVKNSAHRLPLHKLLEHPWIVQNA 424
```

Search completed: June 20, 2005, 20:04:21
Job time : 129.995 secs